AN ECOLOGICAL VALIDITY STUDY OF EXECUTIVE FUNCTION MEASURES IN CHILDREN WITH AND WITHOUT ATTENTION DEFICIT HYPERACTIVITY DISORDER

By

CHRISTOPHER W. LOFTIS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004
Copyright 2004

by

Christopher W. Loftis
ACKNOWLEDGEMENTS

I would like to thank the members of my dissertation committee, Drs. James H. Johnson, Eileen B. Fennell, Bruce A. Crosson, William M. Perlstein, and Christiana M. Leonard, for their support and assistance on this doctoral dissertation, and for serving as role models during my graduate career on both a professional and personal level.

I would also like to specifically acknowledge the following people for providing mentorship and guidance during my doctoral studies. In following my path toward completion of the doctorate I have learned many lessons that have served to make me wiser and to deepen my appreciation for psychology and all that it has to offer the world.

To Dr. Jim Johnson, I thank you for your generosity, patience, and commitment to me during my graduate career. I thank you especially for your modesty and giving nature, the many thought-provoking conversations, the detailed comments and suggestions on this and other work, and your support and advice on many topics of life.

To Dr. Mark Mahone, I thank you for your wisdom and good cheer, and for showing me that it’s possible to be true to yourself. It is not often that one finds an advisor and colleague that always finds the time for listening to the little problems and roadblocks that unavoidably crop up in the course of academics and life.

To Dr. Eileen Fennell, I thank you for your strong spirit, warm heart, and your commitment to students.
To Dr. Bob Frank, I thank you for your inspiration, encouragement, sound advice, strong work ethic, good teaching and supervision, and good company.

To Dr. Cynthia Belar, I thank you for your generosity in support, counsel, and wisdom over the years since beginning my graduate career, and the encouragement to follow the winding path rather than the express toll road straight to the destination.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................ iii

TABLE OF CONTENTS............................................................................................................. v

LIST OF TABLES................................................................................................................ vii

LIST OF FIGURES................................................................................................................ viii

ABSTRACT.............................................................................................................................. ix

CHAPTER

1 INTRODUCTION ................................................................................................................ 1

Attention Deficit Hyperactivity Disorder ............................................................................ 2
Executive Function ............................................................................................................. 8
Assessment of ADHD and Executive Function................................................................. 12
Current Study.................................................................................................................. 22
Hypotheses....................................................................................................................... 27

2 METHODS ....................................................................................................................... 29

Participants and Procedures............................................................................................. 29
Participants ...................................................................................................................... 29
Recruitment .................................................................................................................... 30
Eligibility ........................................................................................................................ 30
Data Collection ............................................................................................................... 32
Measures ........................................................................................................................ 34
Structured Interview ...................................................................................................... 34
Behavior Rating Scales ................................................................................................... 34
Intelligence and academic achievement screening measures ........................................... 37
Standardized Neurobehavioral EF Measures ................................................................. 38
Unstandardized Neurobehavioral EF Measures ............................................................... 46
Statistical Analyses.......................................................................................................... 51
Determination of Significance and Effect Size ................................................................. 51
Analyses of Covariance .................................................................................................. 52
Correlation ....................................................................................................................... 53
Multitrait-Multimethod Matrix ........................................................................... 54

3 RESULTS ........................................................................................................ 60

Sample Characteristics ...................................................................................... 60
  Sex Effects ...................................................................................................... 60
  Age Effects ..................................................................................................... 60
  IQ Effects ....................................................................................................... 62

Tests of Study Hypotheses .............................................................................. 63
  Group Comparisons ........................................................................................ 63
  Multitrait Multimethod Analyses ................................................................... 68
  Evaluation of convergent and divergent validity ............................................. 70

Additional Analyses ........................................................................................ 77

4 DISCUSSION ................................................................................................... 79

APPENDIX

A AGE CORRELATIONS .................................................................................... 97
B IQ CORRELATIONS ......................................................................................... 98
C CORRELATIONS AMONG ALL MEASURES ............................................... 99

LIST OF REFERENCES ...................................................................................... 100

BIOGRAPHICAL SKETCH ............................................................................... 117
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Age and FSIQ Means by Group and Sex</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>Measures of Working Memory and Inhibitory Control</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>Subject Demographics, Reading Achievement, and Severity of ADHD Core Symptoms</td>
<td>61</td>
</tr>
<tr>
<td>4</td>
<td>Group Performance on the BRIEF</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>EF Performance of ADHD-C/HI, ADHD-I, and Control groups</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>EF Performance of ADHD and Control Groups</td>
<td>67</td>
</tr>
<tr>
<td>7</td>
<td>Correlations of BRIEF and EF performance measures</td>
<td>69</td>
</tr>
<tr>
<td>8</td>
<td>Comparisons of EF correlations with IQ</td>
<td>71</td>
</tr>
<tr>
<td>9</td>
<td>Comparisons of Working Memory and Inhibitory Control measures</td>
<td>73</td>
</tr>
<tr>
<td>10</td>
<td>Average Correlations for Inhibitory Control Measures</td>
<td>76</td>
</tr>
<tr>
<td>11</td>
<td>Average Correlations for Working Memory measures</td>
<td>76</td>
</tr>
<tr>
<td>12</td>
<td>Correlations between the BRIEF and CPRS</td>
<td>77</td>
</tr>
<tr>
<td>13</td>
<td>Comparisons of the BRIEF and CPRS</td>
<td>78</td>
</tr>
<tr>
<td>14</td>
<td>Correlations of Age with Rating Scales and Performance Measures</td>
<td>97</td>
</tr>
<tr>
<td>15</td>
<td>Correlations of IQ with Rating Scales and Performance Measures</td>
<td>98</td>
</tr>
<tr>
<td>16</td>
<td>Correlations of BRIEF and Performance Measures</td>
<td>99</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Schematic of behavioral paradigm for the simple Go No-Go task.</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>Sample Multitrait-Multimethod Matrix</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>Simplified Multitrait-Multimethod Matrix</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>Structure of Proposed WM and IC Multitrait-Multimethod Matrix</td>
<td>59</td>
</tr>
<tr>
<td>5</td>
<td>Multitrait-Multimethod Matrix of Working Memory and Inhibitory Control</td>
<td>74</td>
</tr>
</tbody>
</table>
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

AN ECOLOGICAL VALIDITY STUDY OF EXECUTIVE FUNCTION MEASURES IN CHILDREN WITH AND WITHOUT ATTENTION DEFICIT HYPERACTIVITY DISORDER.

By

Christopher W. Loftis

August, 2004

Chair: James H. Johnson
Major Department: Clinical and Health Psychology

This study examined the ecological utility of performance-based measures of executive function (EF) in children with Attention Deficit Hyperactivity Disorder (ADHD). Executive function is broadly defined as the ability to initiate, plan, and sustain behavior in the context of goals, rules, or feedback. Executive dysfunction can adversely impact social, academic, and cognitive functioning; and can complicate the identification of comorbid disorders and intervention procedures for children with special health care needs.

The divergent and convergent validity of working memory (WM) and inhibitory control (IC) were evaluated through the use of behavior rating scales and several performance measures of EF. The predictive validity of EF measures in differentiating between children with and without ADHD was explored with particular emphasis on the ability of WM and IC measures to differentiate among children with ADHD. Measures of EF in this study included behavior ratings from parents (Behavior Rating Inventory of
Executive Function) and performance measures (Delis-Kaplan Executive Function Scale and additional experimental performance-based EF measures).

Analyses of covariance controlling for age and IQ revealed group differences between children with and without ADHD on rating-scales measures of EF, but inconsistent group differences on performance measures of EF. Behavior ratings of WM and IC also better differentiated among children with different ADHD subtypes than did performance measures. Performance measures that placed relatively greater demands on WM were most likely to detect EF impairments identified by parents on the rating scales, whereas performance measures of IC differentiated groups better.

Correlations demonstrating convergent validity of WM and IC across all measures were modest at best; however, there was a pattern of correlations within the behavior rating scales and performance measures to confirm that WM and IC are dissociable as independent constructs. In addition, the association of WM with symptoms of inattention and of IC with symptoms of hyperactivity-impulsivity provided strong evidence for divergent validity and further support that EF is a valid and useful framework for understanding the behavior characteristics of ADHD.
CHAPTER 1
INTRODUCTION

During the past 40 years, research on Attention Deficit Hyperactivity Disorder (ADHD) has traditionally involved study of three core attributes of the disorder: sustained attention, impulsivity, and hyperactivity (American Psychiatric Association, 1994). More recently, investigators have argued that ADHD is also a disorder involving impaired executive function (EF), with some investigators emphasizing behavioral regulation and inhibition components (Barkley, 1997; Tannock, 1998) and others emphasizing motor intention and working memory (Denckla, 1996b; Heilman et al., 1991; Pennington and Ozonoff, 1996). Many cortical structures believed to be critical for EF, particularly the frontal-striatal networks and monoamine systems, have been implicated in the etiology of ADHD (Voeller, 1991). Several investigators have suggested that a model of executive dysfunction better corresponds with neurobiological models of ADHD and has greater explanatory value than the classification of symptoms in current DSM criteria, with some advocating that an adequate understanding of ADHD must include assessment of executive functions in the diagnosis and intervention of children with ADHD children (Goldstein and Goldstein, 1998).

Summarizing across many lines of research in cognitive neuroscience and behavioral neurology, EF can be broadly defined as the ability to initiate, plan, and sustain goal-directed behavior. Varying definitions of EF and disagreement about which aspects of EF are impaired in ADHD children has hindered efforts to develop ecologically valid measures for clinical assessment. In recent years, a handful of
standardized EF batteries and rating scales have become available. This study used two new measures to compare symptom and performance-based assessments of EF in children with ADHD: the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000) and the Delis-Kaplan Executive Function Scale (D-KEFS; Delis, Kaplan, & Kramer, 2001) which provides direct performance measures of the child’s functioning. Convergent and discriminate validity were evaluated through correlations, and additional analyses were undertaken to explore the predictive validity of these tests in identifying subtypes of ADHD.

**Attention Deficit Hyperactivity Disorder**

ADHD is the most common neurodevelopmental disorder in children, with estimates ranging from 3 to 5% (NIH Consensus Statement, 1998). The current edition of the DSM classifies ADHD into three subtypes: Predominantly Inattentive Type, Predominantly Hyperactive-Impulsive Type, and Combined Type (American Psychiatric Association, 1994). This differentiation of symptom clusters was an improvement from DSM III-R, which combined all three symptoms into one diagnostic category. The diagnosis of ADHD requires a persistent pattern of inattention and/or hyperactivity-impulsivity that is developmentally inappropriate with some aspects of the symptom pattern existing before age 7. Impairments must be present in at least two settings, and there must be clear evidence of interference with developmentally appropriate social, academic, or occupational functioning.

Although the three symptom clusters of inattention, impulsivity, and hyperactivity have been considered symptoms of the disorder (either formally or informally) since the turn of the century, the degree to which each cluster has been emphasized has shifted over the decades, in part because researchers have struggled with the heterogeneity of
symptoms and impairments in children with ADHD (Tannock, 1998). Symptoms of inattention can be manifested through careless errors, difficulty persisting on tasks, appearing not to listen to others, frequently shifting from one incomplete activity to another, and failing to complete homework assignments. Children with attention problems can appear easily distracted, and lack habituation to novel or irrelevant stimuli. Symptoms of hyperactivity may be best characterized by excessive amounts of motor activity (such as inability to remain seated in classrooms, on car trips, etc). Adolescents may express hyperactivity through reports of restlessness and dislike of sedentary activities. Symptoms of impulsivity can include impatience, inability to delay gratification or inhibit behavioral responses, inappropriate social behaviors, and a tendency to interrupt others or to not wait their turn.

At least 20 different descriptive terms have been used to identify the constellation of symptoms featured in the current criteria for ADHD, including minimal brain dysfunction, hyperactive child syndrome, hyperkinetic reaction of childhood (DSM II), and attention deficit disorder (ADD). The third edition of the DSM established the label of Attention Deficit Disorder, and greatly expanded the definition of the disorder by allowing diagnosis of subtypes with either maladaptive levels of inattention, impulsivity, and motor activity (ADD with hyperactivity); or attention deficits and impulsivity only (ADD without hyperactivity). DSM-III-R dropped these two subtypes, however, and defined a single category of ADHD that incorporated many of the symptoms found in DSM III. Many studies indicated that the symptoms of ADHD were not unitary as codified by the DSM-III-R definition of ADHD, but the three dimensions of DSM-III symptoms (inattention, impulsivity, and hyperactivity) were not supported either (Lahey,
Pelham, & Schaughency, 1988). Rather, data suggested that two dimensions of symptoms underlie ADHD: one reflecting inattention and the other comprising both hyperactivity and impulsivity. Accordingly, DSM-IV divided the symptoms into two core dimensions: inattention and hyperactive-impulsive; and distinguished three subtypes: those who exhibit maladaptive levels of both dimensions (Combined Type), inattention only (Predominantly Inattentive Type), and hyperactivity-impulsivity only (Predominantly Hyperactive-Impulsive Type). This two-dimensional structure of ADHD has since been supported by both confirmatory factor analyses and discriminant validity studies (Burns, Walsh, & Owen, 1997), but some investigators have begun to make a case for poor impulse control as the core deficit of ADHD (Barkley, 1998, Quay, 1997), arguing that difficulties with attention are secondary to problems with impulsivity. In particular, motor hyperactivity could be seen as the result of inadequate inhibitory control of motor output (mediated by prefrontal and subcortical areas), while inattention might be viewed as inability to inhibit attention to distractors, which as been referred to as interference sensitivity (Teeter and Semrud-Clikeman, 1997). While DSM-III, DSM-III Revised, and DSM-IV differed in classification of symptoms, they are consistent in including two core features: disorganized-inattentive symptoms and hyperactive-impulsive symptoms.

Empirical research suggests that inattention is a cluster separate from impulsivity and hyperactivity, with the latter two symptoms clustering together (Bauermeister, Alegria, Bird, & Rubio-Stipec, 1992). Some have argued that these two clusters may be distinct in etiology, clinical outcome, incidence of comorbidity, and response to treatment (Cantwell and Baker, 1992; Driscoll and Zecker, 1991; Edelbrock, Costello, & Kessler, 1984; Frank and Ben-Nun, 1988; Lahey, Schaugheney, Hynd, & Carlson, 1987;
Kuperman, Johnson, Arndt, & Lindgreen, 1996). Others (Lahey, Schaughency, Strauss, & Frame, 1984) have suggested that children experiencing attention problems without hyperactivity may present twice as frequently in epidemiological studies, while ADD children with hyperactivity may present more frequently to mental health clinics, presumably due to greater severity in disruptive behaviors. Children with the Predominantly Inattentive Type of ADHD may appear more socially withdrawn (Pelham & Milich, 1981) and are more likely to have learning disabilities and academic problems (Edelbrock et al., 1984; Hynd, Lorys, Semrud-Clikeman, & Nieves, 1991; Lahey et al., 1984, 1987). In contrast, children with symptoms of hyperactivity and impulsivity are at greater risk for conduct problems (Edelbrock et al., 1984, Hynd et al., 1991; King & Young, 1982), are more likely to be described as socially unpopular and difficult (Barkley, Fischer, Edelbrock & Smallish, 1990; Cantwell & Baker, 1992; Hynd et al., 1991; Lahey et al., 1984), and may appear less anxious (Lahey et al., 1984, 1987). Overt symptoms of hyperactivity and impulsivity symptoms tend to decline with age (Hart, Lahey, Loeber, Applegate, Green, & Frick, 1995).

Studies have demonstrated impaired performance of ADHD children and adolescents on neuropsychological measures of attention (Douglas, 1988; Fischer, Barkley, Fletcher, & Smallish, 1990; Seidman, Biederman, Faraone, & Weber, 1997), executive functioning (Chelune, Ferguson, Koon, & Dickey, 1986; Fischer et al., 1990; Seidman et al., 1997; Shue and Douglas, 1992), inhibitory control (Fischer et al., 1990; Iaboni, Douglas, & Baker, 1995; Ostererlaan & Seargeant, 1998), and verbal learning and memory (Loge, Staton, & Beatty, 1990; Seidman, Biederman, Faraone, & Milberger, 1995; Seidman et al., 1997; Tannock, Purvis & Schacter, 1993). Investigators have also
identified distinct neuropsychological differences among subtypes of ADHD. Hyperactive children have been found to have longer reaction times than inattentive or control children (Hynd et al., 1988; Sergeant & Scholten, 1985). Goodyear and Hynd (1992) identified a pattern of slower cognitive tempo and greater difficulty with learning new information in children with attention difficulties. DuPaul, Barkley & McMurray (1994) reported that predominantly inattentive children were more off-task during vigilance tasks, performed more poorly on the Coding subtest of the WISC-R, and had greater problems with verbal memory than hyperactive children. Barkely (1994) and Denckla (1994) suggested that children with ADHD Combined Type show poor sustained attention, with improved response to novel stimulation; while those with ADHD inattentive type have more difficulty with focused attention as a result of cognitive slowing or dysfunction.

Trommer, Hoeppner, Lorber, and Armstrong (1988) found differences among ADHD subtypes in impulse control. ADD children with and without hyperactivity were similarly impulsive in the initial phase of a go no-go paradigm. However, ADD children without hyperactivity showed significant improvement as training continued; while ADD children with hyperactivity continued to have problems with impulsivity. Using a sensorimotor soft sign battery with limited cognitive involvement (e.g., eye movement, finger sequencing, choreiform movements), Hern and Hynd (1992) found that there were no differences between normal controls and children with hyperactivity, while children with both hyperactivity and attention deficits exhibited more soft signs than normals at all ages and more soft signs than children with attention deficits only (no hyperactivity) above the age of 8 years. Frank and Ben-Nun (1988) found that hyperactive and attention
disordered children were more impaired on a variety of neurological measures, including abnormal finger sequencing, dyskinesia, choreiform movements, and abnormal hopping. In addition, they were more likely to have had perinatal or neonatal abnormality (50% versus 12%). Hyperactive children also showed greater problems with visual perception, visual sequential memory, and writing performance.

In a review of 22 studies of frontal-lobe functioning in children with attention and hyperactivity difficulties, Barkley, Grodzinsky, & DuPaul (1992) concluded that findings across studies suggest that inattentive, but not hyperactive, children have problems with perceptual motor speed and processing. Further, summarizing the literature, Teeter and Semrud-Clikeman (1997) concluded that ADHD children without hyperactivity may have more difficulty with information processing components of attention, speed of cognitive processing, mental preoccupation, extreme sensitivity to environmental interference, and inwardly directed attention problems, whereas children with ADHD, Combined Type, have more problems with behavioral organization and disinhibition. While some have suggested that neuropsychological differences among subtypes are not well supported (Eiraldi, Power, & Nezu, 1997), it is important to note that many of these studies do not adequately control for comorbid conditions such as conduct/oppositional-defiant disorder, learning disability, or anxiety and mood disorders (Goldstein and Goldstein, 1998).

Less well studied, however, is the relationship between behavior-rating measures and performance-based measures of executive function in ADHD, and whether these measures can be meaningful and useful for the purposes of clinical intervention and prediction of outcomes. At present, diagnostic criteria do not adequately account for the
variability in symptoms throughout development, between males and females, and within subtypes. Goldstein and Goldstein (1998) suggest that while DSM criteria serve as the first step in identifying children with problems with inattention and hyperactivity-impulsivity, it is incumbent on practitioners and researchers to understand the impact of the disorder on development, social adjustment, and cognitive functioning across settings and throughout the lifespan.

**Executive Function**

Various models of executive function (EF) and the neuroanatomical structures that support EF have been widely debated (Barkley, 2000; Denckla, 1994; Fuster, 1989; Goldman-Rakic, 1987; Levin et al., 1991; Stuss, 1986; Welsh & Pennington, 1988), with varying degrees of overlap and consensus as to the overall nature of executive function and specific subdomains. Executive function has become an umbrella term for complex cognitive processes that appear to be intimately associated with the prefrontal cortex (PFC). Most investigators concur that EF coordinates several interdependent cognitive processes that are necessary for purposeful, goal-directed activity (Lezak, 1993, p. 42-43). Definitions of EF commonly encompass cognitive processes responsible for synthesis of external and internal milieu, formation of goals and strategies, maintenance of action and cognitive plans in memory until implemented, and inhibition of other behavior or nonrelevant stimuli. In general, EF can be characterized as critical for carrying out the “how” and “when” to perform behavioral routines (Heilman, 1994), which can be dissociated from knowing “what” to do (Denckla, 1996a).

Luria (1966) stressed the importance of the frontal-lobes in the mediation of activating and regulatory functions. He divided these functions into problem-solving behaviors via posterior regions for processing information (receiving and storing), and
anterior regions for regulation, programming, and verification of mental activity. Luria also stressed the importance of language in the regulation of behavior through the use of either internal linguistic functions or external verbal instructions.

Fuster (1999) defined EF as essential for the planning of behavior that is future-and goal-oriented, and for linking events that are disparate in time. Fuster conceptualizes the entire PFC as a substrate of motor memory that is ideally suited to perform integrative functions where the highest plans and schemas of behavior are formed and represented. His model includes three essential processes: working memory, preparatory set, and inhibitory control. These three constructs are elaborated below to provide a foundation for the subsequent discussion of EF in ADHD.

For Fuster, working memory (which he sees as synonymous with short-term memory) is consistent with non-human primate models of Goldman-Rakic (1987) and Passingham (1995). Working memory is composed of sensory and/or motor representations of environmental stimuli or reactivated perceptual memory, motor memory, or sensory memory. In contrast, preparatory set (alternatively named short-term motor memory or motor attention) is attention directed to action in preparation. Both working memory and preparatory set are commonly impaired by lesions to the dorsolateral prefrontal cortex (DLPFC) and work in tandem to mediate cross-temporal contingencies (meaning that these systems work together to reconcile past learning history with the future-oriented behaviors). However, Fuster stresses that these processes are not localized in one part of the PFC, as they depend heavily on connections of DLPFC with the medial prefrontal cortex and posterior cortices. Disorders resulting from DLFPC dysfunction are characterized by inattention and a failure to focus. Similar to
Fuster’s model, Goldman-Rakic’s (1987) body of work provides evidence that various prefrontal areas serve to guide future behavior by maintaining representations on-line. She identified prefrontal areas that are unique in the nature of the representation that they serve to maintain, with the function of each prefrontal area determined by its relatively unique connections to posterior and subcortical regions.

Inhibitory control, the third component in Fuster’s model, involves suppression of internal and external influences that may interfere with the temporal sequence of behavior. Inhibitory control can be impaired by lesions to the orbital frontal cortex (OFC), which result in a failure to suppress or inhibit distracting stimuli or memories (sensory or motor) from conflicting with on-going behavioral programs. Disorders of the OFC are characterized by emotional lability, impulsivity, and disinhibition of instincts and sexual behaviors. However, disorders of disinhibition can also result from damage to PFC, subcortical structures, and connections between subcortical structures and the PFC (Alexander, DeLong, & Strick 1986; Lhermitte, Pillon, & Serdaru, 1996).

Various models of primate PFC functioning (Fuster, 1989; Goldman-Rakic, 1987; Luria, 1966) converge to paint a picture of PFC as sensitive to tasks that require planning or programming for future actions, maintaining behavioral plans until executed, and inhibition of irrelevant external and internal milieu. In establishing a common definition of EF, it is important to acknowledge that the diversity of cortical connections from PFC to other cortices and among areas of the PFC enables multiple levels of cognition to interact with one another. Consequently, EF is intimately linked with many other cognitive processes; and performance on neuropsychological tests of EF can be sensitive to impairments in subdomains of cognitions, such as tests of memory, perception, and
language. Burgess, Alderman, Evans, Emslie & Wilson (1997) suggest that validation of
the EF construct depends on the degree to which individuals can be differentially
impaired on aspects of EF in comparison to intact functioning on neuropsychological
measures of intelligence, memory, perception, and so on.

Denckla (1994) advocates a move away from hierarchical models of EF (i.e.,
describing EF as superordinate or higher order processes), and argues for
conceptualization of EF as a “central” control process that coordinates cognitive and
behavioral activities. Her conceptualization results in a strong overlap of EF, attention,
and memory, which is useful from a developmental perspective. In fact, she views the
psychological development from childhood to adulthood to be the development of
executive control. She, along with Fuster and Barkley, insist that language abilities are
integral to many EF activities because of the role of language in mediating rule-governed
behaviors.

Several studies indicate that EF skills develop in a stagelike fashion during
childhood (Passler, Isaac, & Hynd, 1985; Welsh, Pennington, & Groisser, 1991), and that
improvements in EF tasks during childhood coincide with growth spurts in frontal-lobe
development (Bell & Fox, 1992; Levin et al., 1991; Thatcher, 1991, 1992; Welsh &
and Anderson (1988) indicate that EF skills undergo significant development between the
ages of 3 and 12 years, especially in the years between 6 and 12. Several studies have
found that many aspects of EF are at or near adult levels of performance by age 10 to 12
(Appelof and Augustine, 1986; Chelune and Baer, 1986; Passler et al., 1985). There is
evidence that aspects of planning and memory rehearsal begin to emerge as early as
infancy (Bruner, 1973). In a study of 100 children age 3 to 12, Welsh et al. (1991) found that EF emerged in three stages of skill integration and maturation: at age 6, age 10, and during adolescence. Organized strategic and planning behaviors were detected as early as age 6. Adult-like performance on more complex measures of organized search ability with greater hypothesis testing and impulse control was evident by age 10. Tasks of verbal fluency, motor sequencing, and complex planning skills were not at adult levels at age 12, suggesting that these skills continue to develop through adolescence and young adulthood. It is important to note that measures used for assessment of EF in adults may not maintain adequate validity when task complexity is reduced for use with children and adolescents. Even young adults demonstrate substantial variability on traditional tests of EF such as the Category Test, the Trail Making Test, and the Wisconsin Card Sorting Task (Shute and Huertas, 1990; Tranel, Anderson, & Benton., 1994).

Assessment of ADHD and Executive Function

Several lines of evidence link ADHD to poor executive control and dysfunction of the frontal-subcortical systems. It has been suggested that tests which are more heavily dependent on PFC systems are more likely than other tests to reveal deficits between ADHD and normal children (Chelune, Ferguson, Koon, & Dickey, 1986). Deficits in sustained attention, inhibitory control, organization, and motivation have been observed in children and adults with damage to the prefrontal circuits (Tranel et al., 1994; Stuss and Benson, 1986). Primate and neuropsychological studies have implicated the PFC in impaired performance on tasks of planning or working memory (Fuster, 1999, Goldman-Rakic, 1987) and in the modulation of basal ganglia circuits (Alexander et al., 1986). Neuroanatomical and lesion evidence reveal that the frontal-lobes have a protracted course of development (Fuster, 1999; Thatcher, 1991) and that proficiency on traditional
EF tasks improves through late adolescence (Tranel et al., 1994; Welsh et al., 1991). Casey, Castellanos, Giedd & Marsh (1997) suggest that the right PFC has a role in suppressing attentional and behavioral responses to salient, but irrelevant distractors, while basal ganglia play a role in the execution of these behavioral responses.

Current theories posit that symptoms of ADHD result from dysfunction of the cortico-striatal-thalamo-cortical network due to structural abnormalities of PFC, basal ganglia, and cerebellum; or result from abnormalities in dopaminergic and noradrenergic pathways originating in brainstem nuclei (locus ceruleus, substantia nigra, and ventral tegmental area). Investigators have proposed a variety of prefrontal cortex models for ADHD, including frontal-lobe dysfunction (Castellanos, Giedd, Hamburger, & Marsh 1996; Hynd, Semrud, Lorys, & Novey 1990), delayed frontal maturation (Chelune et al., 1986), and dysfunctional subcortical-frontal motor systems (Castellanos, Giedd, Eckburg, & Marsh 1994; Giedd, Castellanos, Casey, & Kozuch, 1994). In a review of neuroimaging findings, Hendren, De Backer, & Pandina (2000) found that abnormalities in basal ganglia have been consistently observed in children with symptoms of ADHD, especially those with impulsivity (Casey et al., 1997; Lou, Henriksen, Bruhn, & Borner, 1989). In addition, many MRI studies have reported decreased caudate volumes and reversed asymmetries, although the laterality of these differences and direction of asymmetry have not been consistent across studies (Castellanos, 1997; Hynd, Hern, Novey, & Eliopoulos, 1993). Reduced frontal volume and glucose metabolism has also been identified, but results have been inconsistent as to whether frontal-lobes are reduced bilaterally, or in the right or left hemispheres selectively. Some have also suggested that the behavioral symptoms of ADHD stem from right frontal-lobe and basal ganglia

Voeller (1991) proposed three different types of brain dysfunction that might account for the symptom clusters of ADHD in children, although the neuroanatomical substrates are likely not as separate and isolated from one another as the categorical symptom clusters. First, she describes a “posterior group” with sensory-attentional deficits and hypoarousal that respond to low-dose stimulants. Then she sites evidence for an “anterior group,” with dysfunction of frontal-striatal loops that result in motor intention and inhibitory control deficits. These children may require higher doses of stimulant medication to achieve behavioral control. Finally, a “ventral group” may be characterized by dysfunction in the limbic nucleus-accumbens systems, including the mesolimbic dopamine systems. These children would exhibit high levels of restlessness and locomotor activity, and would respond more variably to stimulant medication.

Models of inhibitory control are featured in many current theories of ADHD (Barkley, 1998; Quay, 1997; Schachar, Tannock, & Logan, 1993; Sonuga-Barke, Stevenson, Thompson, & Lamparelli, 1995). Pennington and Ozonoff, (1996) reviewed findings of executive dysfunction in four developmental psychopathologies (ADHD, autism, conduct disorder, and Tourettes) and found EF to be consistently impaired in ADHD and autism, with specific impairments in inhibitory control unique to ADHD. Tannock (1998) argues that measures of response inhibition offer the best avenue of experimental investigation in children with ADHD. She asserts that models of working
memory in ADHD have been sparse and vague, and that investigations of attention have
offered little specificity or consistency regarding which attentional components are
impaired in ADHD. Denckla (1996) agrees, suggesting that research on output or motor-
affiliated aspects of EF provide greater clarity across the various models and theoretical
formulations of EF. Barkley’s (1997) influential model, based partly upon an
evolutionary model of inhibitory control developed by Bronowski (1977), argues that
disinhibition and poor self-regulation are the core deficits of ADHD; and that other
executive functions are dependent on inhibitory control. Specifically, working memory
(verbal and nonverbal), self-regulation, and reconstitution (a process of synthesizing and
analyzing cognition to generate novel, complex, and purposeful behaviors) are all
dependent on the integrity of inhibitory control.

It is often the goal of developmental neuropsychology studies to identify the
primary and core neurocognitive deficits of disorders. Many researchers have noted
similarities between children with ADHD and patients with EF deficits or prefrontal
cortex lesions (Benton, 1991; Grattan & Eslinger, 1991; Heilman et al., 1991), and
studies of children with ADHD have consistently demonstrated deficits in one or more
domains of EF, including sustained attention, working memory, inhibitory control, and
planning (Harris, Schuerholz, Singer, & Reader 1995; Grodinsky & Diamond, 1992;
Reader, Harris, Schuerholz, & Denckla, 1994).

In a review of 18 studies of EF in children with ADHD, Pennington and Ozonoff
(1996) reported that 15 of 18 studies found a significant difference between children with
and without ADHD on one or more EF measures, with ADHD children performing worse
on 40 of the 60 (67%) EF measures. Measures that were most sensitive to differentiating
between groups (based upon average effect sizes and consistency of findings across the studies) included the Tower of Hanoi, Stroop, Matching Familiar Figures Test errors, and Trailmaking Test – Part B. Purer measures of inhibitory control of motor output (such as the GoNo-Go, Anti-saccade, Conflict Motor Task, and NEPSY Inhibition) also consistently detected group differences, as did two measures of working memory (Sequential Memory Task and Self-Ordered Pointing Task). ADHD children also performed more poorly on measures of vigilance and perceptual speed (6 of 8 measures). In addition to findings of EF deficits across studies, Pennington and Ozonoff (1996) identified consistencies in measures that did not differentiate between ADHD children and controls. In particular, their review of studies failed to find significant differences on 10 of 13 measures of verbal memory tests, and 15 of 19 measures of visual-spatial ability.

While Pennington and Ozonoff (1996) provided sufficient evidence that EF measures differentiate between controls and ADHD children, their review also illustrated that there is considerable variability across studies. Furthermore, research suggests that while abnormal scores on performance measures of EF tests may be sensitive to the presence of ADHD, they may lack specificity. In a study of ADHD adults, Lovejoy, Ball, Keats, Stutts, Spain, Janda, & Janusz (1999) demonstrated high positive predictive power in classifying subtypes of ADHD with several measures of EF, but also evidenced problems with a high number of false negatives. Grodzinsky and Barkley (1999) found good positive predictive power for seven EF tests (ranging from 80 to 90%) administered to children aged 6 to 11 years, but rates of negative predictive power were modest (from 50 to 66%). Sensitivity was also poor (ranging from 5 to 43%), as were the levels of false negatives (averaging 40%). Dykman and Ackerman (1991, 1993) have argued that while
findings from performance-based measures may be statistically associated with diagnoses obtained from observer rating measures, they do not by themselves reliably differentiate between subtypes of ADHD.

The lack of consistent results across studies may be attributed to several factors, including differences in theoretical definitions and assessment of EF (Barkley, 1997; Denckla, 1996a), variability in medication and dosing effects (Nigg, Hinshaw, & Halperin, 1996; Pliszka, Carlson, & Swanson, 1999, Solanto; 1998), variability in diagnostic criteria (Tannock, 1998; NIH, 1998), and differences in degree of organization and structure in the testing situation (Osmon, 1999); as well as differences in EF performance due to the effects of age (Anderson, 2002), sex (Houghton et al., 1999; Rucklidge & Tannock 2002; Siedman, 1997;), and intelligence (Ardilla, Pineda, & Rosselli, 1999; Arffa, Lovell, Podell, & Goldberg, 1998; Dodrill, 1997, 1999).

In his review of ecological validity of neuropsychological tests, Sbordone (1996) asserted that although neuropsychological tests are can helpful in identifying differences among clinical groups and normals, they were often poor at predicting behavior outside of the clinic or neuropsychological laboratory. He concluded that although neuropsychological tests provide beneficial information under some circumstances (e.g., tests of aphasia), these tests should not be used in isolation and without consideration of their relevance to predict functioning in everyday situations. This is particularly important when one considers that executive functioning may be inadequately assessed within a structured assessment setting, given that the examiner provides organization, guidance, cueing, sets time limits, and controls transitions between tasks. The testing environment may not be sufficient to tap core EF deficits in planning, inhibitory control,
and flexibility, which may be more clearly evident in day-to-day contexts such as the home and school. Many have noted that children with ADHD can perform adequately during clinical assessments (Baron, 2003). Furthermore, inconsistencies between performance and behavior rating measures of EF have been well documented (Eslinger & Damasio, 1985). Individuals who behave appropriately in the clinic or obtain normal performance on EF measures in a structured testing situation, may be rated to have significant problems with social judgment and self-regulation when faced with similar situations in real-life (Mesulam, 1985) and have great difficulty making decisions in real life (Stuss & Buckle, 1992).

Given that use of neuropsychological measures in predicting real-world EF behaviors is modest at best, it is important to supplement test data with information obtained from observation of the patient during testing and with information obtained from others who interact with the patient in a variety of other situations. Sbordone and Guilmette (1999) recommended that information about the patient’s behavior be obtained from several sources (medical, academic, and vocational records; significant others; etc.) in order to be compared with neuropsychological data. If neuropsychological data are consistent with functional data, then it is likely that the test data are ecologically valid. In contrast, if the neuropsychological data are inconsistent with functional data, then the ecological validity of the test data are suspect; and further investigation of the data should be undertaken before arriving at a final diagnosis or prediction about the patient’s abilities.

Gioia and colleagues (2001) reported that behavior-rating measures of EF are sensitive to detecting executive functions in a variety of populations, including high
functioning autism and Asperger’s Syndrome (Gilotty et al., 2002), Pervasive Developmental Disorders (Kenworthy, Guy, & Wallace, 2000), brain injuries (Mangeot et al., 2002; Nichols et al. 2002), epilepsy (Slick, Sherman, Connolly, & Eyrl, 2002), low birth weight (Taylor, 2000), Tourette’s Syndrome (Coscia, Ris, Huth, & Gilbert, 2002) and Acute Lymphoblastic Leukemia (Mautz et al., 2002). Gioia et al. (2002) also demonstrated that behavior ratings identify EF profiles of different clinical groups that are consistent with expectations from the literature. In a study comparing EF profiles of children with ADHD, Combined Type (ADHD-C), ADHD, Predominantly Inattentive Type (ADHD-I), moderate traumatic brain injury (TBI), severe TBI, Autistic Spectrum Disorders (ASD), and reading disorders, Gioia et al. (2002) found that children with ADHD and ASD exhibited greater problems overall than children with reading disorders or severe TBI, who were in turn more impaired than children with moderate TBI. Furthermore, the profile of elevations on rating scales was different across the groups. Ratings for children with ADHD-C and ASD were elevated across all scales, while children with ADHD-I and reading disorders exhibited greater difficulties on metacognitive aspects of EF (such as working memory, planning, organization, and self-monitoring), but were not impaired on inhibitory control, shifting, and regulating emotions. Children with ADHD-I were also more severely impaired on the metacognitive scales than were children with reading disorders. While children with ADHD-I had deficits in inhibitory and emotional control, they were not as severe or consistent as the deficits in ADHD-C. Children with ASD demonstrated greater difficulty with shifting than did any of the other clinical groups.
Another issue that further complicates assessment of EF is the obvious fact that EF tasks necessarily enlist the whole brain, and not just one region. Implicit in this notion is the subsequent confound that many EF tasks depend on the integrity of other cognitive process (visual perception, language, procedural knowledge, etc.) (Osmon, 1999). Therefore it is important to assess integrity of more basic cognitive functions when assessing EF, particularly in children (Baron, 2003). Denckla (1996b) argues that content or domain-specific competencies must be adequate before general control processes can evolve. She recommends assessment of EF through coupling of measures that can differentially control for the content processes attributed to language, memory, attention, and so on. For example, when assessing executive abilities involving the verbal domain, one could compare performance on the Vocabulary subtest of the Wechsler Intelligence Scale for Children, a theoretically less EF-demanding measure of stored lexicon, with performance on letter fluency, a measure that makes relatively greater demands on EF processes. Using this method, Denckla (1996b) found that the difference between z-scores on these tasks was greater for ADHD children than for normal controls. In addition, the magnitude of the difference between z scores was not as great when the WISC Vocabulary subtest was compared to the Boston Naming Test, another measure of reasonably low EF load, suggesting that the differences in the Vocabulary-letter fluency comparison was successful in tapping executive control aspects of the verbal domain. Thus, when assessing performance on EF tasks, it is also important to assess functioning on subdomains of cognitive functioning.

Many studies have demonstrated that EF can be reliably differentiated from subdomains of cognitive functioning and fractionated into distinct EF constructs in adults
(Burgess and Shallice, 1994, 1996; Damasio, 1996; Duncan et al., 1995; Owen, Sahakian, Hodges, & Summers, 1995; Robbins, 1996; Robbins et al., 1995; Stuss, Shallice, Alexander, & Picton, 1995; Wilson, Evans, Emslie, Alderman, & Burgess, 1998); however, findings have been less consistent in children (Anderson, 2002, Pennington and Ozonoff, 1996; Welsh, Pennington, & Groisser, 1991). This is not surprising considering that maturation of prefrontal cortex and development of EF skills continue through late adolescence and into adulthood (Thatcher, 1991, 1997). While research has clearly demonstrated that aspects of EF are present across childhood (Levin et al., 1991, Welsh et al., 1991), the developmental trajectory of these processes are not precisely understood, nor are the interactive effects of EF skills at different stages of development (Anderson, 2002). In addition, one has to consider that many studies of EF in children have involved downward extensions of EF tasks used in adults, which in effect make adult performance the marker of frontal-lobe development and EF ability. It is only recently that that developmental psychologists and neuropsychologists have begun to develop unique paradigms for assessing the developmental continuum of EF (Welsh and Pennington, 1988).

Level of intellectual ability may also confound performance on measures of EF by imposing limits on scores or restriction of range on test performance (Dodrill, 1997). For example, Mahone, Hagelthorn, Cutting, Schuerholz, Pelletier, Rawlins, & Denckla (1999) demonstrated differences between ADHD children and controls on commission errors and response variability, but only for children with average IQ and not for those with high or superior IQ. Reader, Harris, Schuerholz, and Denkla (1994) have found that ADHD children with above average IQ were impaired on a number of EF tasks, and in
particular, significant linear relationships were identified on a selected subset of content-
matched tests that differed in EF task demands.

While some have found correlations between IQ and EF (Arffa, Lovell, Podell, &
Goldberg, 1998; Boone, 1999), a review of the literature by Welsh and colleagues (1991)
asserted that many EF measures are not correlated with IQ, including Motor Planning
(Golden, 1981), Verbal Fluency (Ardilla, Galeano, & Rosselli, 1998), Wisconsin Card
Sorting Test (Ardila, Pineda, & Rosselli, 2000; Boone, Ghaffarian, & Lesser, 1993;
Riccio, Hall, Morgan, & Hynd, 1994), and Trail Making Test (Ardilla et al., 2000;
Waldemann, Dickson, Monahan, & Kazelskis, 1992). Summarizing across studies,
Goldstein and Goldstein (1998) concluded that less than 10% of the variance in verbal IQ
is accounted for by the presence of ADHD, although he notes that ADHD, and in
particular hyperactivity and impulsivity behaviors, significantly impair the use of
performance and organizational skills (Hinshaw, Lahey, & Hart, 1993, Sonuga-Barke,
Houlberg, & Hall, 1994). Overall, it appears that ADHD children are able to acquire age
appropriate knowledge and skills, but may be impaired in their ability to efficiently and
effectively make use of their knowledge base.

**Current Study**

Executive dysfunction significantly impacts everyday functioning. A person’s
ability to adapt to and navigate their home work and social environments is determined
by the ability to organize oneself, plan and coordinate activities, monitor and self-regulate
behavior, and inhibit maladaptive responses based upon changing conditions in the
environment determine. Despite the wide variety of cognitive tests that assess various
aspects of EF, there has been little work to validate the use of these measures in
predicting real world functioning (Sbordone, Seyranian, & Ruff, 2000), particularly in
children where characterization of EF is less specified. Although many performance-based measures and caregiver behavior checklists exist for assessing a wide range of behaviors and adaptive functioning skills in children, comprehensive measures of EF in children and adolescents are relatively new and have not yet been well explored.

Delis and colleagues (Delis, D., Kaplan, E., & Kramer, J., 2001) recently published a performance-based battery of executive functions for children and adults, the Delis-Kaplan Executive Function Scale (D-KEFS). The D-KEFS contains nine tests that evaluate flexibility of thinking, inhibitory control, problem solving, planning, impulse control, concept formation, abstract thinking, and creativity in both verbal and spatial modalities. Although the subtests on the D-KEFS were derived from existing neuropsychological tests used in adults (e.g. Stroop, Trails Making Test, verbal fluency, etc.), there are several points that support the use of this measure in exploring the ecological validity of children. First, it is the first neuropsychological assessment battery to provide normative data across a variety of common EF tests in a national sample spanning 8 to 89 years of age, which is a significant advantage in exploring the development of EF skills through childhood and adolescence. Second, the authors took care to develop scoring systems that were based upon neurocognitive principles for evaluating the component processes of tasks thought to be especially sensitive to frontal-lobe dysfunction. In addition, the authors incorporated several process measures that may enable better understanding of strengths and weakness of EF abilities.

As discussed above, Gioia and colleagues (2000) developed the Behavior Rating Inventory for Executive Function (BRIEF), a behavior checklist measure, to assess the multidimensional nature of EF in children and adolescents. It is currently available in
teacher, parent, and self-report forms. The BRIEF contains eight scales theoretically and statistically derived to assess domains of EF (Inhibit, Shift, Emotional Control, Initiation, Task Organization/Planning, Environmental Organization, Self-Monitoring, and Working Memory). The published measure consists of two factors: the Behavioral Regulation Index (BRI) that consists of the Inhibit, Shift, and Emotional Control scales, and the Metacognition Index (MCI) consisting of the Initiation, Task Organization/Planning, Environmental Organization, Self-Monitoring, and Working Memory scales. A composite summary, the Global Executive Composite, is also given, although the authors strongly recommend examining differences between the BRI and MCI indices.

Gioia, Isquith, Retzlaff, and Espy (2002) recently identified a three factor structure of the BRIEF based upon factor analysis of data from a diversity of clinical populations that included ADHD, Autism Spectrum Disorder, Tourette Syndrome, affective disorders, and seizure disorders. In this study, the authors examined the underlying structure of the BRIEF by testing four competing factor structures (1 factor, 2 factors, 3 factors, and 4 factors) for model fit and parsimony. Results suggested a three factor structure consisting of Behavioral Regulation, Emotional Regulation, and Metacognition factors, with the Monitor scale of the Metacognition Index from the two factor model subdividing into Self-Monitoring and Task Monitoring Scales. In this model, the Behavioral Regulation factor was comprised of the Inhibit Scale with the newly derived Self-Monitor scale while the Emotional Regulation factor consisted of the Shift and Emotional Control Scales that were previously included in the Behavioral Regulation Index. The Metacognition factor was comprised of the second subdivision of the Monitor items, the Task Monitor Scale, along with the other scales originally comprising this...
index (Initiate, Working Memory, Plan/Organize, and Organization of Materials). The authors stated that this three-factor structure better captured two aspects of regulatory control - inhibition of external behavior and internalized emotional control and flexibility. They suggest that this model is more consistent with Barkley’s (1997) theory that executive function consists of (1) inhibitory control having a unique and separable role; (2) emotional regulation playing an integral role, and (3) reconstitution, or the metacognitive aspects of executive function that involve the ability to synthesize novel, complex behavioral plans.

Although both the D-KEFS and the BRIEF assess a number of EF constructs, this study focused on the constructs of working memory (WM) and inhibitory control (IC) for several reasons. First, given the large number of constructs that have been proposed as distinct processes under the EF umbrella, it seemed necessary to limit the number of variables explored to ensure adequate power for analyses and make interpretation of results manageable. Second, WM and IC are arguably the most prominently and consistently featured constructs of the many models of EF in the literature, and both appear to be critical for the development of EF skills. Cohen and Servan-Schreiber (1992) suggest that task demand for WM and IC are the two critical dimensions of executive function, and that any task requiring both of these is more likely to engage the PFC. Recent computational models from these authors along with Kimberg and Farah (1993) suggest that WM and IC are intrinsically linked so that increased activation of WM processes allow inhibition of competing response alternatives, while inadequate activation of IC mechanisms allows prepotent, but inaccurate responses to dominate behavioral outcomes. Pennington and Ozonoff state, “Tasks that are high in both their
working memory and inhibition demands are more likely to tax the PFC, although tasks that have a very high demand for either alone are also hypothesized to be prefrontal tasks.” Although these processes can be viewed as independent from one another, and evidence suggests that they can be differentially impaired (Gioia et al., 2002, Pennington and Ozonoff, 1997), Goldman-Rakic (1997) suggests that this independence does not mean they need to be mediated by separate cognitive mechanisms.

Thirty-three possible EF constructs were identified by a conference committee from the National Institute of Child Health and Human Development (Eslinger, 1996). Of these, participants were in 40% or greater agreement on the following six constructs as key components of EF: 1) self-regulation, 2) sequencing of behavior, 3) flexibility, 4) response inhibition, 5) planning, and 6) organization of behavior. Notably, all of these would seem to load heavily upon WM and IC, and the conference participants concluded that WM and IC were the constructs most likely to lead to greater specification of EF functioning (Eslinger, 1996).

WM and IC have also been demonstrated to differentiate EF profiles in childhood disorders (Anderson, 2002; Gioia et al., 2002; Pennington & Ozonoff, 1998). With respect to ADHD in particular, the population of study here, Ozonoff and Jensen (1999) identified a double dissociation between ADHD and autism groups such that children with autism showed difficulties with planning and cognitive flexibility but not inhibitory control, whereas children with ADHD showed the reverse pattern. Pennington and Ozonoff (1996) identified a similar pattern across several studies in which children with autism were more severely impaired on tasks of WM (Siegel & Ryan, 1989) while only ADHD children were impaired on tasks of IC (Bennetto et al., 1996). WM and IC have
also been shown to differentiate between ADHD subtypes (Gioia et al., 2001; Mahone et al., 2002; Pennington and Ozonoff, 1998).

**Hypotheses**

Focusing on the constructs of working memory (WM) and inhibitory control (IC), this study had two primary aims. The first was to investigate the predictive validity of EF measures, assessed through self-report and through direct assessment of the child’s performance, in differentiating between children with ADHD and normal controls, and in differentiating between children with subtypes of ADHD. The second was to explore the convergent and divergent validity of the WM and IC constructs by examining relationships between performance and behavior checklist measures of EF. Regarding the second aim, a Multitrait Multimethod Matrix analytic approach was employed to evaluate the convergent and divergent validity of WM and IC. Correlational analyses were used to explore relationships between performance-based measures of EF (D-KEFS, Go No-Go paradigm, and additional measures) with a behavior checklist measure of EF (BRIEF) and with other measures of behavioral functioning. The following hypotheses were proposed:

**Hypothesis #1:** Children with ADHD would show significant impairment on behavior checklist and performance based measures of EF compared to normal controls.

**Hypothesis #2:** Children with all subtypes of ADHD would show significant impairments on behavior checklist and performance based EF measures of WM.

**Hypothesis #3:** Only ADHD children with Combined subtype would show significant deficits on measures of IC assessed by each of these methods.
**Hypothesis #4**: Ratings on the BRIEF would be more strongly correlated with performance-based measures of EF overall than with measures of intellectual functioning.

**Hypothesis #5**: Performance and behavior checklist measures of similar EF constructs would correlate more strongly with each other than with measures of theoretically different aspects of EF. In particular:

a) Performance measures of WM (e.g., D-KEFS Category Switching, D-KEFS Tower Test, WISC-III PI Spatial Span, CANTAB Self-Ordered Pointing, Counting Go No-Go) would correlate more strongly with a behavior checklist measure of working memory (BRIEF) than with measures of other EF constructs assessed using the same method.

b) Performance measures of IC (e.g. D-KEFS Trails, D-KEFS Color-Word Interference, Conflicting Motor Response, Lateral Gaze, and Simple Go No-Go) would correlate more strongly with a behavior checklist measure of inhibition (BRIEF) than with measures of other EF constructs.
Participants and Procedures

Participants

76 children participated and were assigned to three groups: 39 children without ADHD or other neurodevelopmental or learning disorders who served as controls, 28 children diagnosed with ADHD, Combined Type or Hyperactive-Impulsive Type (ADHD-C/HI), and 9 children diagnosed with ADHD, Inattentive type (ADHD-I).

Details of the diagnostic and testing procedures are provided below. All participants had a FSIQ (WISC-III or WISC-IV) of 80 or above (range 81-144), and none met criteria for reading disability. The sample was predominantly Caucasian (77%), male (68%), and right-handed (78%; Edinburgh Handedness Inventory: Oldfield, 1971).

Table 1. Age and FSIQ Means by Group and Sex

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender (n)</th>
<th>Age</th>
<th>FSIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>F (18)</td>
<td>11.22</td>
<td>120.5</td>
</tr>
<tr>
<td></td>
<td>M (21)</td>
<td>11.14</td>
<td>116.05</td>
</tr>
<tr>
<td></td>
<td>Total (39)</td>
<td>11.18</td>
<td>118.1</td>
</tr>
<tr>
<td>ADHD-C/HI</td>
<td>F (8)</td>
<td>10.75</td>
<td>115.86</td>
</tr>
<tr>
<td></td>
<td>M (20)</td>
<td>12.65</td>
<td>108.75</td>
</tr>
<tr>
<td></td>
<td>Total (28)</td>
<td>12.11</td>
<td>110.59</td>
</tr>
<tr>
<td>ADHD-I</td>
<td>F (2)</td>
<td>14.5</td>
<td>103.5</td>
</tr>
<tr>
<td></td>
<td>M (7)</td>
<td>11.57</td>
<td>102.57</td>
</tr>
<tr>
<td></td>
<td>Total (9)</td>
<td>12.22</td>
<td>102.78</td>
</tr>
</tbody>
</table>

Note. ADHD = Attention-Deficit Hyperactivity Disorder; C/HI = Combined/Hyperactive-Impulsive; I = Inattention; F = Female; M = Male; FSIQ = Wechsler Intelligence Scale for Children
Recruitment

Participants were recruited from four clinics within Kennedy Krieger Institute (KKI): Neuropsychology; Developmental Cognitive Neurology; Center for Learning and its Disorders; and, Psychiatry. In addition to the KKI clinics, children with and without ADHD were recruited from local area pediatricians, schools, social/service organizations (e.g., Boy/Girl Scouts), local chapters of Children and Adults with Attention-Deficit/Hyperactivity Disorder (CHADD), and from advertisements in the community (e.g., postings at libraries).

Eligibility

All children entering the study met the following the criteria:

- Age between 8 years – 0 months and 16 years – 11 months
- Full Scale IQ estimate of 80 or higher based upon on a standardized IQ test given either during a prior school assessment (completed within one year of the current assessment) or on the Wechsler Intellectual Scale for Children-Fourth Edition (WISC-IV; Psychological Corporation, 2003) given at time of testing,
- No history of speech/language disorder or a Reading Disability either screened out before a visit, based on prior school assessment (completed within one year of the current assessment) or (if accepted for a screening visit) RD could also be based on a statistically significant discrepancy between a child's WISC-III Full Scale IQ score and his/her Basic Reading subtest score from the Wechsler Individual Achievement Test®— Second Edition (WIAT–II, Psychological Corporation, 2001), or a standard score below 85 on the Basic Reading Subtest, regardless of IQ score, or performance of less than the 25th percentile on the Basic Reading subtest of the WIAT determined by our research team’s screening visit testing results.
- No evidence of deafness, blindness, cerebral palsy, epilepsy, autism, or psychosis. Diagnosis of ADHD was determined by a structured parent interview Diagnostic Interview for Children and Adolescents, Fourth Edition; DICA-IV) and administration of ADHD-specific and broad behavior rating scales (Conners’ Parent Rating Scale-Revised, Long Form, CPRS).

Participants were accepted with those comorbidities found most commonly with ADHD, namely Oppositional Defiant/Conduct Disorder (with ADHD) and symptoms of anxiety; however, children with any other DSM diagnoses were excluded. DSM-IV
criteria were used to evaluate for all three ADHD subtypes: Predominantly Inattentive Type (ADHD-I), Predominantly Hyperactive-Impulsive Type (ADHD-HI), and Combined Type (ADHD-C). Children were assigned to the ADHD-I group if the T-score on the CPRS Scale L (DSM criteria for Predominantly Inattentive Type) was 65 or greater AND their T-score was 60 or less on the CPRS Scale M (DSM criteria for Predominantly Hyperactive-Impulsive Type). Children were diagnosed as ADHD-HI if the T-score on the CPRS Scale M (DSM criteria for Predominantly Hyperactive-Impulsive Type) was 65 or greater AND their T-score was 60 or less on the CPRS Scale L (DSM criteria for Predominantly Inattentive Type). Children were assigned to the ADHD-C group if the T-Score on both Scales L and M were both rated as 65 or higher.

Only 4 children were diagnosed with ADHD, Predominantly Hyperactive-Impulsive Type, presumably due to the close relationship between hyperactive-impulsive behaviors and difficulties with inattention. Consequently, the ADHD-C and ADHD-H/I children were combined into one ADHD-C/HI group. Every attempt was made to match the groups on the basis of age, FSIQ, gender, and minority status.

Controls were additionally required to meet the following criteria: (1) no history of mental health services for behavior or emotional problems, (2) no parent or teacher report of previous diagnosis of ODD or CD; (3) percentiles scores below 1.5 standard deviations above the mean on the ADHD subscales of Conners’ Parent Rating Scale-Revised (CPRS; Conners, 1997), and (4) no history of academic problems requiring school based intervention services or history of defined primary reading or language-based learning disability, as established through medical history, psychological testing, or parental and
teacher interview. All children in the control group had T-Scores below 61 on both the Inattention and Hyperactivity-Impulsivity Scales of the CPRS.

**Data Collection**

Prior to scheduling the appointment, parents of participants were briefly interviewed over the telephone in order to obtain demographic information (name, age, date of birth, parent(s) names, phone number, address), referral source, school and developmental history. If the child was determined to be eligible via the brief screen, the DICA was administered by phone prior to scheduling an appointment at KKI. Parents of children with ADHD were asked to not administer medication on the day before and of testing (for a total of 48 hours before testing).

On the day of the appointment, parents completed questionnaires and rating scales while the child completed testing. The total time period required for both parents and children was approximately 2 ½ hours. Participants provided written consent (caregivers) and assent (children) before beginning testing and received a copy of the consent form. Caregivers completed a brief background questionnaire that provided information regarding the child’s age, date of birth, grade level, handedness, maternal and paternal education level, maternal and paternal occupation, significant medical history (i.e., age of diagnosis, medications), school history, and history of learning disabilities. In addition, caregivers completed the following questionnaires with the specific instruction to complete measures according to the child’s behavior when NOT on medications: the Conners’ Parent Rating Scale-Revised, Long Form (assesses impulsivity and ADHD behaviors) and the Behavior Rating Inventory of Executive Function (assesses executive cognitive functioning).
Parents were also given a packet on the day of testing for the child’s teacher that contained instructions for signing the consent form and completing the Conners’ Teacher Rating Scale-Revised, Long Form and the Behavior Rating Inventory of Executive Function. Teachers were asked to complete the measures according to the typical day-to-day performance of the child and to return materials in a pre-paid, self-addressed envelope. Because results from the DICA and parent rating scales were confidential, teachers were not informed of the child’s diagnosis and the examiner was not able to interview teachers directly. Thus, if a child with ADHD took medication while in school, it was impossible to ascertain whether the teacher was familiar with the child’s behavior off medication. Consequently, a decision was made to not include the teacher data in analyses due to the confounding effect of teachers completing measures blind while parents completed measures according to a child’s behavior off medication and children completed performance testing while not on medication.

During the 2 ½ hour appointment, children were administered the following standardized measures on the day of testing: the Wechsler Intelligence Scale for Children, Fourth Edition if no prior estimate of intelligence was available, the Word Reading subtest from the Wechsler Individual Achievement Test, Second Edition, the Spatial Span subtest of the Wechsler Intelligence Scale for Children, Third Edition, Processing Instrument, and the Color-Word Interference, Tower, Trails, and Verbal Fluency subtests of the Delis-Kaplan Executive Function System. In addition, the following unstandardized neurobehavioral measures were administered to assess WM and IC: Conflicting Motor Response, Lateral Gaze Fixation, CANTAB Self-Ordered Pointing, Simple Go No-Go, and Counting Go No-Go.
Measures

Five types of measures were utilized for this study: a) structured interview, b) behavior rating scales (information provided by caregivers); c) intelligence and academic achievement screening measures; and, d) standardized neuropsychological measures of EF, and e) unstandardized neurobehavioral measures of EF.

Structured Interview

Diagnostic Interview for Children and Adolescents, Fourth Edition (DICA-IV):
The parent version of the DICA-IV (Reich, Welner, & Herjanic, 1997) was administered to parents about their child. This is a semi-structured interview that is designed for determining selected current and retrospective psychiatric diagnoses. The modules administered were those assessing present and retrospective reports of: Attention Deficit Hyperactivity Disorder, Conduct Disorder, Oppositional Defiant Disorder, Major Depressive Disorder, Bipolar Disorders, Dysthymic Disorder, Separation Anxiety Disorder, Panic Disorder, Generalized Anxiety Disorder, Specific Phobia, Obsessive Compulsive Disorder and Adjustment Disorders. The DICA-IV has been reported to be reliable for DSM-IV diagnoses.

Behavior Rating Scales

The child’s parent(s) or guardian(s) completed the following standardized rating scales in order to assist in the diagnosis of ADHD and to obtain data on behavior ratings of EF.

Behavior Rating Inventory of Executive Function (BRIEF): The BRIEF (Gioia et al., 2000) Parent Form consists of 86 items derived from theoretical and empirically based definitions of EF and from items submitted by practicing clinical neuropsychologists. Children were evaluated on a 3-point Likert scale (never, sometimes,
often). The BRIEF contains three general indices: Behavioral Regulation (consisting of three scales: Inhibit, Shift, and Emotional Control), Metacognitive problem solving (consisting of the five remaining scales: Initiation, Task Organization/Planning, Environmental Organization, Self-Monitoring, Working Memory), and a Global Executive Composite. Higher ratings are indicative of greater perceived impairment. The BRIEF Parent Form was normed on 1,419 children age 5 to 18. Factor analytic studies of the normative sample support the existence of two underlying factors, which were used to develop the MCI and BRI indices. The BRIEF scales have demonstrated strong psychometric properties: internal consistency, stability over short periods of time, and interrater (teacher-parent) agreement in the appropriate range. Mean internal consistency ratings ranged from 0.82 to 0.98. Three-week test-retest correlation for clinical populations ranged from 0.72 to 0.84 for the three indices and 8 scales.

Of the eight non-overlapping BRIEF scales, the diagnostic scales of Working Memory and Inhibit along with the BRIEF Indices (BRI and MCI) were the primary dependent measures in this study. While the remaining six (i.e., Initiate, Plan/Organize, Organization of Materials, Monitor, Emotional Control, and Shift) may be of relevance to EF and to studies of children with ADHD, this study focused on working memory and inhibitory control given their prominence in models of EF and their strong theoretical relationship to ADHD symptoms.

The Working Memory scale correlates with a variety of attention scales, such as those included within the BASC (Reynolds & Kamphaus, 1994), CBCL/TRF (Achenbach, 1991), Conners’ Rating Scales (Conners, 1989), and the ADHD Rating Scale IV (DuPaul, Power, Anastopoulos & Reid, 1998), providing evidence for
convergent validity. Examples of items from the Working Memory Scale include: “forgets what he/she was doing,” or “has trouble remembering things, even for a few minutes.” The Inhibit scale correlates strongly with measures of restlessness, impulsivity, overactivity, behavior problems, and aggression, and demonstrates secondary correlations with attention problem scales. Items from the Inhibit Scale include: “talks at the wrong times,” and “gets out of control more than friends.”

Both the Working Memory and Inhibit scales correlated moderately with scales reflecting social difficulties, consistent with the observation that children who have attention and, in particular, impulse control problems, also have social difficulties. Equally important are the low correlations with a variety of scales that reflect behavioral and emotional difficulties that should not be related to inattention, impulsivity, and hyperactivity. The pattern of low correlations with scales measuring somatic complaints, anxiety, and depression provide evidence of divergent validity for the Working Memory and Inhibit scales.

Conners Parent Rating Scale (CPRS): The CPRS (Conners, 1997) is a broadband child behavior rating scale that can be completed by parents of children aged 3 to 17 years. The CPRS items required the parent to rate behaviors on a 3-step response scale (not true, sometimes true, always true). The CPRS was originally normed on 11,000 community-based children and adolescents throughout the United States and Canada. The revised version contains items that match the DSM-IV criteria for ADHD. Normative data for the current revised scales come from the ratings of approximately 2,000 parents and teachers. One-week test-retest reliability for the CPRS subscales averaged 0.68. The CPRS has demonstrated appropriate convergent, discriminant, and
divergent validity. The Scales L (DSM-IV: Inattentive) and M (DSM-IV: Hyperactive-Impulsive) of the CPRS were used to assign children to the three groups and as dependent variables in correlational analyses.

**Intelligence and academic achievement screening measures**

Wechsler Intellectual Scale for Children-Fourth Edition (WISC-IV): The WISC-IV (Psychological Corporation, 2003) is a commonly used and well-normed measure (2,200 children who were divided into 11 age groups with 100 boys and 100 girls at each age level) individually administered for assessment of intellectual functioning of children aged 6 years through 16 years, 11 months. The WISC-IV consists of 15 subtests that are summarized in five composites reported in standard scores (Full Scale IQ, Verbal Comprehension, Processing Speed, Working Memory, Perceptual Reasoning). Reliability coefficients for the WISC-IV composite scores ranged from .99 (Processing Speed) to .97 (Full Scale IQ), which were identical to or slightly better than the corresponding composite scores on the Wechsler Intellectual Scale for Children-third edition (WISC-III). Reliability coefficients for the subtests ranged from .70 (Cancellation) to .90 (Letter-Number sequencing). The mean retest scores for all subtests were higher than the mean test scores from the first administration, with effect sizes ranging from .08 (Comprehension) to .60 (Picture Completion). In general, the test-retest gains were less pronounced for the verbal subtests than for other subtests. The Full Scale IQ (FSIQ) was used to screen for adequate intellectual functioning (above FSIQ of 80) and as a dependent measure for correlational analyses.

processes involved in many of the WISC–III subtests. Reliability coefficients for Spatial Span forward (.66) and backward (.75) were in the moderately high range. The mean retest scores for all subtests were higher than the mean test scores from the first administration, with stability quotients in the moderate range. Spatial Span Backwards was used the dependent measure for this test as a measure of nonverbal working memory.

Word Reading subtest of the Wechsler Individual Achievement Test®—Second Edition (WIAT–II): The WIAT-II Word Reading (Psychological Corporation, 2001) test assessed early reading (phonological awareness) and word recognition and decoding skills. The WIAT was developed to be used in conjunction with the WISC and provides age (sample of 2,950) and grade-based (sample of 3,600) Standard Scores for age 4 years through college level students. Reliability coefficients for Word Reading were high (.97), as were stability coefficients (.98). This measure was used to screen for reading disability in all children according to the criteria described in the eligibility section due to the potential developmental effects of reading disability on performance measures of IC and working memory (Rucklidge JJ & Tannock R, 2002)

**Standardized Neurobehavioral EF Measures**

The Delis-Kaplan Executive Function System (D-KEFS): The D-KEFS (Dean C. Delis, Edith Kaplan, Joel H. Kramer, 2001) assessed key components of executive functions with nine tests that evaluate flexibility of thinking, inhibition, problem solving, planning, impulse control, concept formation, abstract thinking, and creativity in both verbal and spatial modalities. Designed to be interesting and engaging for examinees, the D–KEFS' game-like format encourages optimal performance without providing "right/wrong" feedback that can create frustration in some children and adults. It is the first comprehensive battery of EF with a national normative sample (1,750 children and
adults). The sample was divided into 16 age groups consisting of separate norms for ages 8 through 15, then group normative data for the ages 16-19 and every decade thereafter (e.g., 20-29, 30-39, etc.). The sample contained 75 children in each of the 8, 9, 10, and 11-year-old age groups, and 100 children for each of the 12, 13, 14, and 15-year-old age groups, with all age groups relatively equal in gender and race.

The D-KEFS does not provide composite scores across the tests because “each D-KEFS test was developed as an individual, stand-alone instrument” (D-KEFS technical manual, page 53). Limited validity data were provided in the D-KEFS technical manual for tests due to its recent release and because several measures are “modifications of long-standing clinical or experimental tests”. The validity of D–KEFS instruments that are modified tests (i.e., the Stroop procedure, Trail Making Test, verbal and design fluency tests, tower tasks, the twenty-questions procedure, and proverb interpretations) has been demonstrated in numerous neuropsychological studies conducted over the past 50 years or more” (D-KEFS technical manual, page 47). Consequently, the following tests from the D-KEFS were selected for this study based upon their demonstrated utility in the literature to assess EF in child populations: Trail Making Test, Verbal Fluency, Color-Word Interference, and Tower Test. Psychometric properties and rational for choosing each of these measures are described below.

D-KEFS Trail Making Test: This test consisted of a visual cancellation task and a series of connect-the-circle tasks. The primary EF task is Condition 4: Number–Letter Switching, which provided a means of assessing flexibility of thinking on a visual–motor sequencing task. The other four conditions of this test allow the examiner to quantify and derive normative data for several key component processes necessary for performing the
switching task (e.g., visual scanning, number sequencing, letter sequencing, and motor speed). Internal consistency coefficients for D-KEFS Trail Making Test ranged from moderate (0.57 at age 10) to high (0.79 at age 14). Most reliability coefficients were in the moderate range with the highest test-retest correlation observed for the motor speed task.

The D-KEFS Trail Making Test is adapted from the Army Individual Test Battery Trail Making Test (TMT; Army Individual Test Battery, 1944). The TMT has two subtests: Part A, which requires sequencing of numbers, and Part B, which requires sequencing of letters and numbers (similar to the D-KEFS Trails Condition 4). The TMT is one of the most widely used neuropsychological tests because of its sensitivity to general brain dysfunction (Reitan, 1971) and to impairment in a variety of cognitive domains (Lezak, 1995). In addition to the ability of the D-KEFS to assess key component processes involved in performing the switching task, an important difference between the D-KEFS and TMT tests is that the D-KEFS Trail Making Test requires scanning of the stimuli across a 17 by 11 inch sheet of paper while the TMT version is limited to a single 8 ½ by 11 sheet of paper.

Performance on TMT Part B has been shown to improve during the childhood years (Anderson, 1998) and several studies have identified reduced performance on the TMT Part B in developmental disorders and generalized cerebral dysfunction in children (Barkley & Grodzinsky, 1994; Beers et al., 1994). TMT Part B has also been shown to discriminate between ADHD children and normal controls (Gorenstein, Mammato, & Sandy, 1989; Dykman and Ackerman, 1991; Shue and Douglass, 1992) with an average effect size of .75 (Pennington and Ozonoff, 1988). Kelly (2000) found that improved
scores on TMT B from 7 to 13 years of age were associated with improved IC. Thus, the dependent measure from this test was the Trail Making Test Condition 4 Total Time as a measure of IC.

D-KEFS Verbal Fluency Test: This test is comprised of three testing conditions: Letter Fluency, Category Fluency, and Category Switching. For the Letter Fluency condition, the examinee is asked to generate words that begin with a particular letter as quickly as possible. Answers are not allowed to be names of people, places, or numbers, or repeats of the same word with different endings. For the Category Fluency condition, the examinee is asked to generate words that belong to a designated semantic category as quickly as possible. The last condition, Category Switching, requires the examinee to generate words, alternating between two different semantic categories as quickly as possible. For each trial of each condition, the examinee is allowed 60 seconds. This test measured the examinee’s ability to generate words fluently in an effortful, phonemic format (Letter Fluency), from overlearned concepts (Category Fluency), and while simultaneously shifting between overlearned concepts (Category Switching).

The Letter Fluency condition yielded the highest internal consistency coefficients, which ranged from moderate to high (.68 at age 8 to .81 at age 13). The internal consistencies were lower for Category Fluency (.53 at age 15 to .72 at age 12) and Category Switching (.37 at age 8 to .62 at age 11-13). The pattern of test–retest correlations was consistent with those reported for the internal consistency values. The test–retest measures for both the Letter Fluency and Category Fluency conditions had moderate to high reliability. Category Switching had lower correlations than those for Letter Fluency and Category Fluency. In the retest sample, performance improved on
average from the first testing to the second testing. The Category Switching total score had the lowest interval score change.

Verbal fluency measures have a long history in neuropsychological assessment (Lezak, 1995). Most versions of verbal fluency include a letter fluency task that assess the ability to generate words in response to a letter cue, and a semantic fluency task that assess the ability to produce words in response to a category cue. Letter fluency is associated with anterior cerebral functioning in adults while semantic fluency is associated with more posterior regions. Letter fluency is considered more difficult than semantic fluency for children (Halperin, Healey, Zeitchik, & Ludman, 1989), but there is some disagreement as to when children attain adult levels of performance, with some suggesting that adult levels are reached around age 10 (Regard, Strauss & Knapp, 1982), while others suggest that performance continues to improve into adolescence (Welsh et al, 1991). There are also inconsistent findings regarding whether fluency tests differentiate between children with and without ADHD. Pennington and Ozonoff (1988) identified a small effect size of 0.27 for studies that found ADHD children to perform worse on letter fluency tasks (Loge et al., 1990; Grodzinsky and Diamond, 1992), while others have failed to identify differences (Mahone, Koth, Cutting, Singer, & Denckla, 2001; McGee Williams, & Moffitt, 1989). Verbal Fluency performance has been shown to be associated with intelligence and vocabulary (Tager-Flusberg, 1985).

Baron (2003) stated that verbal fluency measures consist of two components (1) a linguistic component associated with left hemisphere function, and (2) an executive function component associated with the frontal lobes. She noted that there is a working memory component due to the need to retrieve words while remembering previously
stated words, and an inhibitory component due to the need to self-monitor and avoid rule-breaks. A significant change on the D-KEFS Verbal Fluency Test from previous versions of verbal fluency was that the rules for the D-KEFS version (i.e., No names of people, places, or numbers, and no repetition of the same word with different endings) are displayed on a cue card during the letter fluency trials, thus providing assistance for utilizing the rules and possibly providing external support for inhibitory control of rule breaks. While many studies have chosen the total number of words generated during letter fluency as the variable of interest, the D-KEFS Letter Fluency test was not chosen for analyses in this study for the reasons discussed above: the uncertainty of age effects, the poor and inconsistent ability of letter fluency measures to differentiate between ADHD and control groups, the association with IQ, and the possible reduced demands placed upon inhibitory control due to displaying the rules in the D-KEFS version. Instead, Category Switching, the third condition of the D-KEFS Verbal Fluency Test, was chosen as the dependent variable for this study as a measure of working memory.

D-KEFS Color-word Interference Test: The primary executive function measured from this test is the examinee’s ability to inhibit an overlearned verbal response (i.e., reading the printed words) in order to generate a conflicting response of naming the dissonant ink colors in which the words are printed. This test has two baseline conditions that measured key component skills of the higher-level tasks: basic naming of color patches (Condition 1) and basic reading of color-words printed in black ink (Condition 2). The third test, Condition 3: Inhibition, is the traditional Stroop task, for which the examinee must inhibit reading the words in order to name the dissonant ink colors in which those words are printed. The D–KEFS test also included a new EF task, Condition
4: Inhibition/Switching. For this condition, the examinee is asked to switch back and forth between naming the dissonant ink colors and reading the words. This condition is thus a means of evaluating both inhibitory control and cognitive flexibility. Internal consistency coefficients for this test range from moderate (.62 at age 13) to high (.79 at age 8). Reliability coefficients were in the moderate to high range for the key color–word variables. Scores on the second testing were generally higher, suggesting improved performance after being exposed to the test.

The Color-word Interference Test is based upon the Stroop Color Word Test, which also has a long history in psychology as a measure of IC (MacLeod, 1991). It has shown to be sensitive to brain damage (Golden, 1976; Dodrill, 1978) and to differentiate between ADHD and normal control groups (Gorenstein et al., 1989; Boucuganani and Jones, 1989; Grodzinsky and Diamond, 1992) with an average effect size of 0.69. The dependent measure used for analyses with this test was the Condition 3: Inhibition Total Time to Completion as a measure of IC.

D-KEFS Tower Test: This test assessed several key executive functions, including spatial planning, rule learning, inhibitory control of impulsive and perseverative responding, and the ability to establish and maintain the instructional set. The objective of this task was to move disks, varying in size from small to large, across three pegs to build a designated tower in the fewest number of moves possible. In constructing the target towers, the examinee must follow two rules: (a) move only one disk at a time and (b) never place a larger disk over a smaller disk. Each trial began by the examiner presenting the disks on the pegs in a predetermined starting position and displaying a picture that showed the ending position of the disks, that is, the tower to be built. A Total
Achievement score was obtained that was the sum of achievement scores, including bonus points, for all items administered. The raw score could range from 0 (all items failed) to 30 (i.e., all towers correctly built within the time limit and with the minimum number of moves possible).

The Total Achievement Score had a moderate (.43 at age 14) to high (.84 at age 10) coefficients across age groups. For most age groups, the reliability was in the moderate range. The test–retest correlations were in the moderate range. Average performance improved from the first testing to the second testing. Although validity data for the D-KEFS Tower are limited, the Tower of Hanoi and Tower of London, from which the D-KEFS Tower was adapted, have been frequently used in investigations of EF in children. They are considered to be appropriate for adaptation to pediatric populations because the tasks are generally challenging and attractive to children, and they incorporate a range of difficulty levels so that even young children are able to complete the first few items (Anderson, 1998). Performance on the Tower tests have been shown to improve with age and to differentiate between clinical groups (Anderson, 2002; Baron, 2003; Levin et al, 1990; Welsh et al., 1991) with an average effect size of 1.08 (Pennington and Ozonoff, 1988). In a study of normal child and undergraduate controls, Baker, Segalowitz, and Ferlisi (2001) compared the two most common scoring procedures for the Tower Tests, total accuracy versus speed of completion. Results indicated that scoring for accuracy showed continued development into adulthood whereas scoring for speed reached adult levels by 13 years of age. Consequently, the dependent measure used for analyses with this test was the Total Achievement Score as a measure of working memory.
Unstandardized Neurobehavioral EF Measures

The following unstandardized neurobehavioral measures were included to provide additional method variance in the assessment of WM and IC. All of these measures were borrowed directly from or based upon existing measures in the literature that have been argued to assess either WM or IC. Because the D-KEFS measures are new and have not been as well validated as other versions upon which the D-KEFS tests are based, it was important to include additional measures in order to insure adequate assessment for the construct validity analyses in this study. In addition, it was likely that few of the D-KEFS measures were “pure” measures of either WM or IC alone. These additional, relatively simpler, measures strengthened the ability of the study to assess WM and IC through performance tests. In their review, Pennington and Ozonoff (1998) noted that purer measures of IC of motor output consistently identified group differences between ADHD children and controls. Three of the measures they identified were used for this study, a Go No-Go paradigm, an anti-saccade task (Lateral Gaze), and the Conflicting Motor Response task. These authors also identified the Self-Ordered Pointing task from the Cambridge Automated Neuropsychological Test and Battery as a measure of WM that has shown utility in assessing deficits in children with ADHD. The Self-Ordered Pointing task was included in this study along with an additional measure of WM, the Counting Go No-Go paradigm that is an unstandardized neurobehavioral measure similar in appearance to the Simple Go No-Go paradigm but designed to increase WM demands.

Simple Go No-Go: This task was developed as a measure of motor response inhibition for fMRI studies in order to minimize extraneous cognitive and behavioral variables (Mostofsky et al., 2003). Cues consisted of green (Go) and red (No-Go)
spaceships (see Figure 1 below). Subjects are instructed to push a button as fast as they can to a green spaceship and to refrain from pushing when they see a red spaceship. Use of familiar color elements (green for “Go”; red for “No-Go”) contributes to relative isolation of skeletomotor IC on this task by minimizing superimposed cognitive (WM) processes. The task was presented on a PC computer running E-Prime software. Cues appear on the screen for 300 msec and are presented once every 1.8 seconds (1.5 sec interstimulus interval). Cues are weighted towards Go cues at a ratio of 3:1 (75% Go cues; 25% No-Go cues), intensifying the need to inhibit a rapid, habitual skeletomotor response. The total time of the task was 10 minutes.

It has been argued that this task most directly assesses the construct of IC because it requires an all-or-none decision about action or nonaction (compared with cognitive forms of IC, such as interference control) (Rubia, Taylor, & Smith., 2001). Effect sizes for Go No-Go paradigms in the literature with ADHD children range from 0.53 to 0.78 (Pennington and Ozonoff, 1998). FMRI findings from the Mostofsky et al, 2003 paper using this task are discussed below following the description of the Counting Go No-Go task; however, it is worth noting here that No-Go events (i.e., correct inhibitory control of a motor responses to red spaceships) was associated with signal change in the pre-SMA. For this study, commission errors were used as the dependent measure for assessment of IC.

Counting Go No-Go: The design of the Counting Go No-Go task (Mostofsky, 2003) was similar to that of the simple Go No-Go task except that subjects were instructed to push a button (with their right index finger) as quickly as possible in response to green spaceships AND to red spaceships preceded by an even number of
green spaceships but not for red ships following an odd number of green ships. The number of green spaceships before a red spaceship ranged from three to five, and the task was balanced for the odd and even sets of red spaceships. A simple practice involving three separate scenarios—an even trial, an odd trial, and one of each—was completed by each participant prior to scanning. The total time of this task was also 10 minutes.

Figure 1. Schematic of behavioral paradigm for the simple Go No-Go task.

Mostofsky et al. (2003) utilized the Simple and Counting Go No-Go tasks to compare changes in activation for tasks with different WM demands – one with a high working memory load and one with low. As described above, the Simple Go No-Go task with low working memory load required subjects to push a button in response to green spaceships but not red spaceships. The Counting Go No-Go task with a high working memory load required subjects to respond to green spaceships as well as to those red
spaceships preceded by an even number of green spaceships. For both tasks, Go responses were associated with signal change in the left primary sensorimotor cortex, supplementary motor area (SMA) proper, and anterior cerebellum (right > left). For the simple task, No-Go events were associated with activation in the pre-SMA, while the working memory-loaded “counting” task elicited additional No-Go activation in the right dorsolateral prefrontal cortex. The findings suggested that neural contributions to response inhibition might be task dependent; the pre-SMA appears necessary for inhibitory control of unwanted movements, while the dorsolateral prefrontal cortex is recruited for tasks involving increased working memory load.

Conflicting Motor Response (Christensen, 1975): Adapted from the Luria-Christensen Battery, the Conflicting Motor Response task (Christensen, 1975) assessed motor response inhibition. Subjects were told, "If I show you my finger, you show me your fist; if I show you my fist, you show me your finger." The examiner, using the left hand, presented each of the two gestures 24 times (for a total of 48 presentations) in a fixed pseudo-random sequence, at a rate of one per second. The subjects were instructed to respond with their dominant hand as quickly as possible. The task therefore required the subject to inhibit the prepotent tendency to mimic the examiner. The variable of interest relevant to response inhibition was total number correct on the 48 trials. ADHD children have been shown to perform poorly on this task relative to normal controls (Shue and Douglass, 1992) and relative to autism subjects (Loftis et al., in preparation), with an average effect size of 1.24.

Lateral Gaze fixation: Adapted from a battery of tasks reported in Kertesz et al (1985), this measure of lateral fixation was analyzed as a measure for motor persistence.
Patients were asked to sustain a lateral gaze for 20 seconds. The examiner held a pencil about 45 degrees from the plane between the examiner and patient midlines in the patient's right visual field. The trial terminated at 20 seconds or earlier if the subject's eyes deviated from the indicated fixation point. The procedure is given for the right visual field, then the left visual field, and then repeated for the right and left visual fields. The dependent measure was the sum of left and right lateral gaze scores, with two trials on right side and two trials on the left side for a total of four trials of twenty seconds each. A perfect score was 80 seconds, indicating no errors in holding the gaze for 20 seconds twice on the right and twice on the left.

Self-Ordered Pointing of the Cambridge Automated Neuropsychological Test and Battery (CANTAB): The CANTAB Self-Ordered Pointing (SOP) task (CeNeS Cognition, 1996) was administered as a measure of WM via a portable Pentium MMX processor-based panel PC with a LCD flat panel display touch screen built into the computer (Advantech model PPC-120T-RT). The screen was placed in front of the subject within arms’ reach. Participants were shown colored squares (boxes) on the computer screen and were instructed to “search through” a number of boxes for a blue token, by touching each one to reveal what is inside. The goal was to collect all the blue tokens and fill up a container on the right side of the screen. Returning to an empty box where a target has already been found was counted as a “between-search error.” After four practice trials with three boxes, the participant was presented with 12 test trials that increased in difficulty by increasing the number of boxes to search. Participants completed four sets of trials with four boxes, then four sets of six boxes, and finally four
sets of eight boxes. The dependent measure for the SOP task was the total between errors for the 4-, 6-, and 8-box problems.

Below in Table 1 is a summary of measures employed in this study, listed by the constructs they were chosen to measure.

<table>
<thead>
<tr>
<th>Measure</th>
<th>WM</th>
<th>IC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIEF</td>
<td>Metacognitive Index</td>
<td>Behavior Regulation Index</td>
</tr>
<tr>
<td></td>
<td>Working Memory Scale</td>
<td>Inhibit Scale</td>
</tr>
<tr>
<td>D-KEFS</td>
<td>Category Fluency Total Switching</td>
<td>Trails 4 Total Time</td>
</tr>
<tr>
<td></td>
<td>Tower Total Achievement</td>
<td>Color Word Interference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Condition 3: Inhibition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total Time</td>
</tr>
<tr>
<td>WISC</td>
<td>Spatial Span Backward Scaled Score</td>
<td></td>
</tr>
<tr>
<td>CANTAB</td>
<td>Self-Ordered Pointing Total Between Errors</td>
<td></td>
</tr>
<tr>
<td>Motor Tasks of Inhibition</td>
<td>Conflicting Motor Response Total Correct</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lateral Gaze Total Time</td>
<td></td>
</tr>
<tr>
<td>Go-No-Go</td>
<td>Counting Omissions</td>
<td>Simple Commissions</td>
</tr>
</tbody>
</table>

**Statistical Analyses**

**Determination of Significance and Effect Size**

Data were analyzed using the Statistical Package for the Social Sciences (SPSS) Windows, version 10.1. A sample size of 31 is suggested to obtain power of .80 in order to detect a difference between scores of one half standard deviation (.5) at an alpha level of .05. Due to the unequal sample sizes of the groups and the multiple correlations employed in this study, significance for this study was determined using a conservative alpha level of 0.01 and a medium effect size (0.5 or greater). Findings from post-hoc
analyses provide p-values near .05 to indicate when results were close to significance.

Effect sizes are reported for correlations and t-tests in Cohen’s d and for analyses of covariance in partial eta squared (\( \eta^2_p \)). Cohen (1988) defined \( d \) as the difference between the means, divided by standard deviation of either group. Cohen argued that the standard deviation of either group could be used when the variances of the two groups are homogeneous; however, in practice, the pooled standard deviation is used (Rosnow and Rosenthal, 1996). For interpreting Cohen’s \( d \), 0.2 is considered a small effect (85% overlap in the distributions of the two groups compared), 0.5 a medium effect (67% overlap), and 0.8 a large effect (53% overlap). The partial eta squared is the proportion of the effect variance plus error variance that is attributable to the effect. For interpreting the partial eta-squared, 0.01 is considered a small effect (accounting for 1% of the variance), 0.06 a medium effect (6% of the variance), and 0.14 a large effect (14% of the variance).

**Analyses of Covariance**

To address Hypotheses 1, 2 and 3, group differences were analyzed with ANCOVA’s controlling for age and IQ to determine if ADHD children differed from control children on behavior checklist and performance measures of EF. It was expected that ADHD children would be significantly more impaired on measures of EF (Hypothesis 1). In addition, children with all subtypes of ADHD were expected to show significant impairments on measures of working memory (Hypothesis 2) while only children with Combined Type of ADHD would show significant problems on measures of inhibitory control (Hypothesis 3). Given the low number of subjects in the ADHD-I
group, data were again analyzed by combining the ADHD-C/HI and ADHD-I groups into one ADHD group, again using ANCOVA’s and controlling for age and IQ.

**Correlation**

To address Hypotheses 4 and 5, Pearson product-moment partial correlations, controlling for age and IQ, were conducted to evaluate whether EF constructs measured by the BRIEF were more strongly correlated with similar constructs measured by performance based EF measures than with IQ measures. Relevant to these analyses, Standards for Educational and Psychological Testing (1999) refers to validity as “validity evidence”, explaining that validity is a multifactorial concept that is evaluated by the degree to which all accumulated evidence supports use of a test to measure a construct. While a test should correlate highly with other measures that assess similar constructs, a test should also not correlate with measures that assess other constructs. The former is referred to as convergent validity, and the latter as divergent validity.

In order to objectively evaluate convergent and divergent validity, it was necessary to test whether correlations between different measures were significantly different from one another by transforming correlations into $z$-scores through Fisher’s $r$ to $z$ transformation and then testing for significant differences through $z$-tests on the transformed correlations (Hays, 1988). Convergent and divergent validity were examined for EF measures as a whole, and then for each of the WM and IC constructs.

To evaluate validity for EF measures, mean correlations of the BRIEF scales with each of the performance measures of EF and with IQ were evaluated to first determine whether the BRIEF as a whole was more strongly correlated with the EF performance measures than with IQ. It was expected that correlations between behavior checklist and
performance based measures of EF would be significantly higher than correlations with measured intelligence (Hypothesis 4).

To evaluate the validity of the WM and IC constructs, correlations were evaluated for the BRIEF Working Memory scale with each of the EF performance measures and correlations of the BRIEF Inhibit scale with each of the EF performance measures. It was expected that correlations between BRIEF Working Memory scale and performance based measures of WM would be significantly higher than correlations between the BRIEF Working Memory scale and performance based measures of IC, and conversely, it was expected that correlations between the BRIEF Inhibit scale and performance based measures of IC would be significantly higher than correlations between the BRIEF Inhibit scale and performance measures of WM (Hypothesis 5).

**Multitrait-Multimethod Matrix**

Results from the correlations were averaged and incorporated into a Multitrait-Multimethod Matrix (MMM; Campbell & Fiske, 1959). A MMM is a simple matrix or table of correlations that is arranged to facilitate evaluation of construct validity across a variety of psychological measures and constructs. This tabular arrangement of correlations between constructs and methods allows detection of variance unique to each construct, variance unique to each method of assessment, and variance that is unique to error variance. The MMM thus allows examination of convergent and divergent validity in conjunction with investigation of variance associated with the method of administration (method variance).
Creating a Multitrait-Multimethod Matrix requires the assessment of two or more constructs by two or more different methods. It is ideal to measure each construct with each measure employed. Below is a model of the MMM (Figure 2; adapted from Trochim, 2001) constructed from a hypothetical study of three constructs assessed by three different methods.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A₁</td>
<td>(0.99)</td>
<td>(0.93)</td>
<td>(0.94)</td>
</tr>
<tr>
<td>B₁</td>
<td>0.51</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>C₁</td>
<td>0.38</td>
<td>0.59</td>
<td>0.67</td>
</tr>
<tr>
<td>A₂</td>
<td>0.57</td>
<td>0.59</td>
<td>0.67</td>
</tr>
<tr>
<td>B₂</td>
<td>0.22</td>
<td>0.58</td>
<td>0.60</td>
</tr>
<tr>
<td>C₂</td>
<td>0.11</td>
<td>0.40</td>
<td>(0.85)</td>
</tr>
<tr>
<td>A₃</td>
<td>0.56</td>
<td>0.43</td>
<td>(0.94)</td>
</tr>
<tr>
<td>B₃</td>
<td>0.23</td>
<td>0.43</td>
<td>(0.92)</td>
</tr>
<tr>
<td>C₃</td>
<td>0.11</td>
<td>0.34</td>
<td>(0.92)</td>
</tr>
</tbody>
</table>

**Figure 2. Sample Multitrait-Multimethod Matrix**

There are several aspects of the matrix that are critical to understanding and interpreting the MMM. First, the matrix consists only of correlations. Second, these correlations can be grouped into three shapes: diagonals, triangles, and blocks. The specific shapes can be referred to as the Reliability Diagonal (top diagonal), Validity
Diagonals (bottom two diagonals), Heterotrait-Monomethod Triangles (3 darker triangles), Heterotrait-Heteromethod Triangles (6 lighter triangles), Monomethod Blocks, and Heteromethod Blocks.

- The Reliability Diagonal coefficients are simply estimates of the reliability of each measure in the matrix. There are as many correlations in the reliability diagonal as there are measures (i.e., in this example, there are nine measures and nine reliabilities).

- The Validity Diagonals are correlations between measures of the same construct assessed by different methods. These values could also be considered monotrait-heteromethod correlations, and are important in evaluating construct validity. Since the MMM is organized into method blocks, there is one validity diagonal in each method block. For example, the correlation between A1-A2 is 0.57. This is the correlation between two methods (1 and 2) measuring the same construct (A). Because the two measures are assessing the same construct, they should be strongly correlated.

- The Heterotrait-Monomethod Triangles are the correlations among measures that share the same method of measurement. For instance, A1-B1 = 0.51 in the upper left heterotrait-monomethod triangle. Strong method variance are indicated if these correlations are high.

- Heterotrait-Heteromethod Triangles are correlations that differ in both construct and method. Generally, because these correlations share neither construct nor method, they should be the lowest in the matrix. For instance, A1-B2 is 0.22 in the example.

- The Monomethod Blocks consist of all of the correlations that share the same method of measurement. There are as many blocks as there are methods of measurement.

- The Heteromethod Blocks consist of all correlations that do not share the same methods.

Campbell and Fiske (1959) proposed a set of rules for evaluating the convergent and divergent validity in a MMM. A simplified version of the MMM is presented below.
Figure 3. Simplified Multitrait-Multimethod Matrix.

The basic principles for interpreting a MMM are:

1. Coefficients in the reliability diagonal should consistently be the highest correlations in the matrix. That is, a construct should be more highly correlated with itself than with anything else. This is uniformly true in the first MMM example with data.

2. Validity coefficients in the diagonals (V) must be greater than zero and sufficiently large to warrant further study. This is essentially evidence of convergent validity.

3. Validity coefficients (V) must be greater than heterotrait-heteromethod validity coefficients (HH, in dotted triangles). For example, in looking at the correlations between Method 1 and Method 2 in the first example, A1-A2 (0.57) is greater than A1-B2 (0.22) and A1-C2 (0.11).

4. Validity coefficients (V) should be higher than all coefficients in the heterotrait-monomethod triangles (HM, in solid triangles), meaning that construct factors should be stronger than methods factors. Note that this is not true in all cases in the first example. For instance, the C1-C2 correlation of .46 is less than heterotrait correlations in the Method 2 heterotrait-monomethod block, suggesting that there might be a methods factor, especially for Method 2.
5. The same pattern of construct interrelationships should be shown in all heterotrait triangles (in both monomethod and heteromethod blocks), indicating that the same pattern of construct interrelationship should be seen in all triangles. The first MMM example meets this criterion very well as the A-B relationship is approximately twice as large as the relationships that involve C in all triangles.

It is important to understand that that interpretation of the MMM requires the researcher to use judgment. It is rare that all of the above principles are met perfectly. Even though some of the principles may be violated, it is possible to still have sufficient evidence for construct validity. The MMM allows researchers to identify weaknesses in measurement while also assessing for construct validity.

A simple matrix using the constructs of WM and IC is presented below (Figure 4). It shows the correlation matrix between measures of WM and IC when assessed by the three different methods used in this study: BRIEF, D-KEFS, and Go No-Go. The convergent and divergent validity coefficients are listed in diagonals outside of the triangles and the reliability coefficients for each method’s assessment of each construct are listed in parentheses along the top diagonal. The table includes correlations between different constructs measured by the same method (in solid triangles; e.g. between Working Memory and Inhibit scales of BRIEF) and between different constructs measured by different methods (dotted triangles; e.g. WM measured by BRIEF, D-KEFS, and Go No-Go). For convergent validity, the validity coefficients (V) should be higher than correlations between different constructs measured by different methods. In addition, they should also be higher than correlations between different constructs measured by the same method (solid triangles). High correlations in the latter, representing method variance, would suggest that scores on the measure are unduly affected by an irrelevant or spurious common factor.
Figure 4. Structure of Proposed WM and IC Multitrait-Multimethod Matrix
CHAPTER 3
RESULTS

Sample Characteristics

Information regarding group differences on the variables of age, sex, IQ, WIAT reading scores, and ADHD core symptoms, indexed by Conner’s Rating Scale indices, are presented in Table 3 below. IQ’s for 20 children were obtained from previous administrations (within one year of testing) of the WISC-III (8 ADHD-C/HI and 12 Controls) while the remaining children were administered the WISC-IV on the day of testing along with the other measures. Comorbid disorders were present in 15 of the children with ADHD-C/HI (13 with Oppositional Defiant Disorder, 2 with Simple Phobia, 1 with Obsessive-Compulsive Disorder and 1 with Generalized Anxiety Disorder) and 3 of the children with ADHD-I were diagnosed with Oppositional Defiant Disorder.

Sex Effects

As can be seen by inspection of Table 3, the three groups were predominantly male, with the percentage of male subjects ranging from 77.78% (ADHD-I) to 53.85 (Controls). There were, however, no significant differences between groups in the ratio of boys and girls across the three diagnostic groups \[ \chi^2 (76) = 3.10; p > .05 \].

Age Effects

There were also no significant between-group differences between the three groups in age [ADHD-C/HI = 12.1; ADHD-I = 12.22; Controls = 11.18: \( F (2, 73) = 2.039; p < .137, \eta^2_p = .053 \)]. It should be noted that when the ADHD-I and ADHD-C/HI groups were
combined into one ADHD group and compared to controls, the age effect became

significant \[ t(74) = -2.028, p = .046 \], with the mean age of the ADHD group being

significantly higher than the mean of the control group. In subsequent analyses with

these measures, age was used as a covariate.

Table 3. Subject Demographics, Reading Achievement, and Severity of ADHD Core

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>ADHD-C/HI</th>
<th>ADHD-I</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age</td>
<td>12.1</td>
<td>2.061</td>
<td>12.22</td>
</tr>
<tr>
<td>% Male</td>
<td>71.42</td>
<td>--</td>
<td>77.78</td>
</tr>
<tr>
<td>WISC FSIQ \text{a,b}</td>
<td>110.53</td>
<td>13.66</td>
<td>102.78</td>
</tr>
<tr>
<td>WIAT Reading</td>
<td>110.57</td>
<td>9.09</td>
<td>111.68</td>
</tr>
<tr>
<td>CPRS - I \text{c}</td>
<td>68.93</td>
<td>11.33</td>
<td>67.63</td>
</tr>
<tr>
<td>CPRS HI \text{d,e}</td>
<td>72.04</td>
<td>12.264</td>
<td>55.5</td>
</tr>
</tbody>
</table>

Note. \text{a} Controls > ADHD-I, \( p < .01 \); \text{b} Controls > ADHD-C/HI, \( p < .05 \); \text{c} ADHD-C/HI, ADHD-I > Control, \( p < .001 \); \text{d} ADHD-C/HI > Control, ADHD-I, \( p < .001 \); \text{e} ADHD-I > Controls, \( p < .05 \); ADHD = Attention-Deficit Hyperactivity Disorder; C/HI = Combined/Hyperactive-Impulsive; I = Inattention; WISC FSIQ = Wechsler Intelligence Scale for Children; WIAT Reading = Wechsler Individual Achievement Test Word Reading, values reported in standard scores; CPRS – I = Conners Parent Rating Scale DSM Inattentive (Scale L), values reported in T scores; CPRS HI = Conners Parent Rating Scale DSM Hyperactive-Impulsive (Scale M), values reported in T scores.

Post hoc analyses assessing the relationship between age and study measures are

presented in Appendix A. Using a conservative alpha level of 0.01 and a medium effect

size as the determination for significance, age was not significantly correlated with WISC

Full Scale IQ, WISC Working Memory Index, WISC Spatial Span, WIAT Reading, or

any of the D-KEFS subtests. This was expected as these measures are standardized for

age. However, correlations for the BRIEF indices approached significance with large

effect sizes. More impaired EF scores were associated with increasing age. Among the

EF measures without norm-referenced scores, there were significant (\( p < .01 \)) correlations

with age for Conflicting Motor Response, Lateral Gaze, Self-ordered Pointing, and

Simple Go No-Go.
IQ Effects

Regarding IQ, there was a significant between-group difference for WISC Full Scale IQ (FSIQ) [ADHD-C/HI = 110.53; ADHD-I = 102.78; Controls = 118.1: $F(2, 74) = 7.976; p< .001, \eta^2_p = .181$]. Here, post-hoc tests (Tukey HSD) revealed the control group to have a significantly higher FSIQ than the ADHD-I group ($p< .01$) and the ADHD-C/HI group ($p< .05$). There were no differences in IQ between ADHD Groups.

Given the presence of age effects on some measures, partial correlations with FSIQ were conducted controlling for age and using raw scores for the BRIEF, CPRS and performance-based EF measures. FSIQ correlations are listed in Appendix B. Using a conservative alpha level of 0.01 and a medium effect size as the determination for significance, the BRIEF Working Memory and Metacognitive Index (MCI), as well as the CPRS DSM IV Inattention scale were significantly correlated with FSIQ in the negative direction, indicating that higher IQ scores were associated with lower ratings (less impaired) on the behavior rating scales. Of the EF performance measures, the WISC Spatial Span Backward, D-KEFS Trail Making Test, D-KEFS Category Switching, D-KEFS Color-word Interference Test Inhibition, and Conflicting Motor Response were also significantly correlated with FSIQ, suggesting that children with higher FSIQ performed more quickly (D-KEFS Trail Making Test), completed more items (WISC Spatial Span Backward, D-KEFS Category Switching), and committed fewer errors (Conflicting Motor Response) on EF performance measures.

The BRIEF Inhibit scale and Behavior Regulation Index (BRI) as well as the CPRS DSM IV Hyperactivity-Impulsivity scales were not significantly correlated (at the .01 level) with FSIQ, nor were the D-KEFS Tower Test, Lateral Gaze, Self-Ordered
Pointing, and the two Go No-Go tasks (Appendix B); however, it should be noted that the effect sizes of many of these measures were in the medium range.

Tests of Study Hypotheses

Group Comparisons

Group differences on the BRIEF: To evaluate whether children with ADHD-C/HI and ADHD-I were rated as more impaired on behavior checklist based measures of EF relative to normal controls (Hypothesis 1), a multivariate analysis of covariance (controlling for age and IQ) was conducted on the BRIEF BRI, MCI, Working Memory scale and Inhibit scale. There was a significant multivariate group effect (Pillai’s) for the four BRIEF scales analyzed \[ F (1,70) = 18.126; \ p < .00001, \ \eta_p^2 = .523 \]. Univariate tests revealed highly significant group differences on the scales predicted to differentiate between children with ADHD and Controls \( p < .0001 \): Inhibit \[ F (2, 72) = 31.067, \ \eta_p^2 = .477 \]; Working Memory, \[ F (2, 72) = 63.102, \ \eta_p^2 = .650 \]; MCI \[ F (2, 72) = 36.930, \ \eta_p^2 = .521 \] and BRI \[ F (2, 72) = 21.679, \ \eta_p^2 = .389 \]. Mean T-scores for the five BRIEF-Parent scales are presented in Table 4 with findings from post-hoc tests indicated where results were significant or close to significance.

Table 4. Group Performance on the BRIEF

<table>
<thead>
<tr>
<th>BRIEF Scales</th>
<th>ADHD-C/HI M</th>
<th>SD</th>
<th>ADHD-I M</th>
<th>SD</th>
<th>Control M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibit Scale (^{b,d})</td>
<td>68.19</td>
<td>12.41</td>
<td>51.63</td>
<td>9.46</td>
<td>47.41</td>
<td>7.97</td>
</tr>
<tr>
<td>Behavioral Regulation Index (^{b,c})</td>
<td>64.07</td>
<td>11.55</td>
<td>56.63</td>
<td>11.34</td>
<td>44.24</td>
<td>9.83</td>
</tr>
<tr>
<td>Working Memory Scale (^a)</td>
<td>68.19</td>
<td>11.06</td>
<td>74.13</td>
<td>5.41</td>
<td>44.16</td>
<td>6.42</td>
</tr>
<tr>
<td>Metacognition Index (^a)</td>
<td>66.15</td>
<td>11.22</td>
<td>69.13</td>
<td>5.79</td>
<td>44.92</td>
<td>8.39</td>
</tr>
</tbody>
</table>

Note. \(^a\) ADHD-C/HI, ADHD-I > Control, \( p < .0001 \); \(^b\) ADHD-C/HI > Controls, \( p < .0001 \); \(^c\) ADHD-C/HI > ADHD-I, \( p < .05 \); \(^d\) ADHD-C/HI > ADHD-I, \( p < .0001 \); All BRIEF scores are mean T scores; BRIEF = Behavior Rating Inventory of Executive Function; ADHD = Attention-Deficit Hyperactivity Disorder; I = Inattention; C/HI = Combined/Hyperactive-Impulsive.
Consistent with Hypotheses 2 and 3, post-hoc tests (LSD) showed that both of the ADHD-I and ADHD-C/HI diagnostic groups were rated higher than the control group on the MCI and Working Memory scales ($p < .0001$), whereas only the ADHD-C/HI group was rated higher on the BRI and Inhibit scales ($p < .0001$) compared to controls. ADHD-C/HI ratings were also significantly higher than ADHD-I ratings on the BRI ($p < .05$) and Inhibit scales ($p < .0001$). There were no significant differences between ADHD-C/HI and ADHD-I on the MCI and Working Memory scales, nor between the control group and ADHD-I on the BRI and Inhibit scales.

Group differences on performance measures of executive function: To evaluate whether children with ADHD-C/HI and ADHD-I were rated as more impaired on performance measures of EF relative to normal controls (Hypothesis 1), a multivariate analysis of covariance (controlling for age and IQ) was used to examine group differences on the 9 performance-based EF measures using raw scores for all measures. The analysis revealed no significant multivariate between group differences $[F (20, 60) = 1.133; p=0.343, \eta_p^2 = .274]$ for the performance-based measures. Group means for the EF performance measures are presented in Table 5 with findings from post-hoc tests indicated where results were significant or close to significance.

Exploratory univariate tests revealed significant between group differences for Conflicting Motor Response $[F (2, 65) = 4.907, p = .01, \eta_p^2 = .131]$ and Lateral Gaze $[F (2, 64) = 11.828, p < .000, \eta_p^2 = .270]$, with post-hoc tests (Tukey HSD) indicating that both ADHD-I and ADHD-C/HI groups performed significantly worse than the control group ($p < .05$) but not from one another. Trends were identified for D-KEFS Tower Total Achievement $[F (2, 68) = 3.070, p = .053, \eta_p^2 = .083]$, Self-Ordered Pointing
Between Errors $F(2, 64) = 2.733, p < .074, \eta^2_p = .092$, and Counting Go No-Go Omissions $F(2, 64) = 2.782, p < .071, \eta^2_p = .095$. Post-hoc analyses (Tukey HSD) revealed that the ADHD-C/HI group performed significantly worse than ADHD-I and controls on Tower ($p = .016$) while ADHD-I performed similarly to the control group. Trends were noted for worse performance of ADHD-C/HI and ADHD-I groups compared to controls on Self-Ordered Pointing ($p = .064$ and $p = .058$, respectively) and Counting Go No-Go ($p = .056$ and $p = .059$, respectively), but the ADHD groups did not perform significantly different from one another.

<table>
<thead>
<tr>
<th>Performance Measures</th>
<th>ADHD-C/HI M</th>
<th>ADHD-C/HI SD</th>
<th>ADHD-I M</th>
<th>ADHD-I SD</th>
<th>Control M</th>
<th>Control SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trails 4</td>
<td>10.41</td>
<td>2.135</td>
<td>8.38</td>
<td>1.847</td>
<td>11.39</td>
<td>2.946</td>
</tr>
<tr>
<td>CF Switching</td>
<td>10.48</td>
<td>2.806</td>
<td>8.63</td>
<td>3.114</td>
<td>11.26</td>
<td>2.892</td>
</tr>
<tr>
<td>CWI Inhibition</td>
<td>10.15</td>
<td>3.29</td>
<td>8.63</td>
<td>3.068</td>
<td>11.66</td>
<td>1.893</td>
</tr>
<tr>
<td>Tower $^b$</td>
<td>10.3</td>
<td>1.75</td>
<td>10.5</td>
<td>2.673</td>
<td>11.68</td>
<td>1.684</td>
</tr>
<tr>
<td>Span Backwards</td>
<td>10.74</td>
<td>2.246</td>
<td>9.75</td>
<td>3.412</td>
<td>11.78</td>
<td>2.551</td>
</tr>
<tr>
<td>Conflicting Motor Resp. $^a$</td>
<td>-0.17</td>
<td>1.123</td>
<td>-0.78</td>
<td>0.687</td>
<td>0.28</td>
<td>0.865</td>
</tr>
<tr>
<td>Lateral Gaze $^a$</td>
<td>-0.32</td>
<td>1.116</td>
<td>-0.79</td>
<td>1.322</td>
<td>0.39</td>
<td>0.606</td>
</tr>
<tr>
<td>SOP Between Errors $^c$</td>
<td>0.26</td>
<td>1.183</td>
<td>0.43</td>
<td>0.742</td>
<td>-0.21</td>
<td>0.899</td>
</tr>
<tr>
<td>Simple Go No-Go</td>
<td>0.00</td>
<td>0.959</td>
<td>0.33</td>
<td>0.931</td>
<td>-0.08</td>
<td>1.054</td>
</tr>
<tr>
<td>Counting Go No-Go $^c$</td>
<td>-0.27</td>
<td>0.896</td>
<td>-0.53</td>
<td>0.477</td>
<td>0.25</td>
<td>1.063</td>
</tr>
</tbody>
</table>

Note. $^a$ ADHD-C/HI and ADHD-I < Control, $p < .05$; $^b$ ADHD-C/HI < ADHD-I and Control, $p < .05$; $^c$ ADHD-C/HI and ADHD-I < Control, $p < .1$; Standardized scores are reported in mean Scaled Scores (mean = 10); Raw scores are reported in z scores (mean = 0). Trails 4 = D-KEFS Trails Condition 4 Total Time; CF Switching = D-KEFS Category Switching Total; CWI Inhibition = D-KEFS Color Word Interference Inhibition; Tower Achievement = D-KEFS Tower Total Achievement; Span Backwards = WISC Spatial Span Backwards; Conflicting Motor = Conflicting Motor Response; Lateral Gaze = Lateral Gaze Total Time; SOP Between Errors = Self-Ordered Pointing Between Errors; Simple Go No-Go = Simple Go No-Go Commissions; Counting Go No-Go = Counting Go No-Go Omissions.

Univariate testing suggested no significant between groups differences on the D-KEFS Trails Condition 4 $F(2, 68) = .705, p = .497, \eta^2_p = .020$, D-KEFS Category Switching Total.
Switching Total \( F(2, 68) = .443, p = .644, \eta^2_p = .013 \), D-KEFS Color Word Interference Inhibition \( F(2, 68) = 2.059, p = .135, \eta^2_p = .056 \), WISC Spatial Span Backwards \( F(2, 68) = .313, p = .732, \eta^2_p = .009 \), and Simple Go No-Go Commissions \( F(2, 59) = 1.358, p = .265, \eta^2_p = .044 \).

Thus, unlike the strong findings from the EF rating scales in support of Hypotheses 1, 2, and 3, results from the EF performance measures weakly supported Hypothesis 1 as children with ADHD were significantly impaired on only 2 of 10 performance measures. Hypotheses 2 and 3 were also not supported since the ADHD-I and ADHD-C/HI groups performed similarly to one another with the exception of one measure, the D-KEFS Tower.

In contrast to the findings above when performance measures were analyzed for group differences between the three groups (ADHD-C/HI, ADHD-I and Control), 5 of the 10 performance measures differentiated between groups at an alpha level of less than 0.05 level when ADHD-I and ADHD-C/HI were combined into a single ADHD group and then compared to the performance of the control group. Covaring for age and IQ, the multivariate analysis of covariance revealed a trend for between group differences \( F(11, 29) = 1.886; p = .084, \eta^2_p = .417 \) for the performance-based measures. This is a much stronger result than when performance was analyzed across the three groups \( p = 0.343, \eta^2_p = .274 \). Group means for ADHD and controls are presented in Table 6 below with findings from post-hoc results indicated where significant or close to significance.

Continuing the use of a conservative alpha level of 0.01, univariate analyses again only revealed significant between groups differences for Conflicting Motor Response \( F(1, 66) = 7.678; p = .007, \eta^2_p = .104 \) and Lateral Gaze \( F(1, 65) = 20.702; p = .000, \eta^2_p = \)
.242]; however, trends with smaller \( p \) values than when analyzed across the three groups were identified for D-KEFS Color Word Interference Inhibition \( [F (1,70) = 3.360; p = .071, \eta_p^2 = .046] \), D-KEFS Tower Total Achievement \( [F (1, 69) = 5.869; p = .018, \eta_p^2 = .078] \), Self-Ordered Pointing \( [F (1,55) = 4.903; p = .031, \eta_p^2 = .082] \), and Counting Go No-Go Omissions \( [F (1,60) = 5.138; p = .027, \eta_p^2 = .087] \).

No group differences were identified for the D-KEFS Trails Condition \( [F (1,70) = .286; p = .595, \eta_p^2 = .004] \), D-KEFS Category Switching \( [F (1,70) = .118; p = .732, \eta_p^2 = .002] \), WISC Spatial Span Backwards \( [F (1,69) = .582; p = .448, \eta_p^2 = .008] \), and Simple Go No-Go Commissions \( [F (1,60) = 2.240; p = .140, \eta_p^2 = .036] \).

<table>
<thead>
<tr>
<th>Performance Measures</th>
<th>ADHD M</th>
<th>ADHD SD</th>
<th>Control M</th>
<th>Control SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trails 4</td>
<td>9.94</td>
<td>2.22</td>
<td>11.39</td>
<td>2.95</td>
</tr>
<tr>
<td>CF Switching</td>
<td>10.06</td>
<td>2.94</td>
<td>11.26</td>
<td>2.89</td>
</tr>
<tr>
<td>CWI Inhibition d</td>
<td>9.80</td>
<td>3.26</td>
<td>11.66</td>
<td>1.89</td>
</tr>
<tr>
<td>Tower a</td>
<td>10.34</td>
<td>1.96</td>
<td>11.68</td>
<td>1.68</td>
</tr>
<tr>
<td>Span Backwards</td>
<td>10.51</td>
<td>2.54</td>
<td>11.78</td>
<td>2.55</td>
</tr>
<tr>
<td>Conflicting Motor Resp. b</td>
<td>-0.32</td>
<td>1.06</td>
<td>0.28</td>
<td>0.87</td>
</tr>
<tr>
<td>Lateral Gaze c</td>
<td>-0.43</td>
<td>1.17</td>
<td>0.39</td>
<td>0.61</td>
</tr>
<tr>
<td>SOP Between Errors a</td>
<td>0.31</td>
<td>1.08</td>
<td>-0.21</td>
<td>0.90</td>
</tr>
<tr>
<td>Simple Go No-Go</td>
<td>0.08</td>
<td>0.95</td>
<td>-0.08</td>
<td>1.05</td>
</tr>
<tr>
<td>Counting Go No-Go a</td>
<td>-0.34</td>
<td>0.81</td>
<td>0.25</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Note. a ADHD < Controls, \( p < .001 \); b ADHD < Control, \( p < .01 \). c ADHD < Control, \( p < .05 \). d ADHD < Controls, \( p < .1 \). Standardized scores are reported in mean Scaled Scores (\( x =10 \)). Raw scores are reported in z scores (\( x = 0 \)). Trails 4 = D-KEFS Trails Condition 4 Total Time; CF Switching = D-KEFS Category Switching Total; CWI Inhibition = D-KEFS Color Word Interference Inhibition; Tower Achievement = D-KEFS Tower Total Achievement; Span Backwards = WISC Spatial Span Backwards; Conflicting Motor = Conflicting Motor Response; Lateral Gaze = Lateral Gaze Total Time; SOP Between Errors = Self-Ordered Pointing Between Errors; Simple Go No-Go = Simple Go No-Go Commissions; Counting Go No-Go = Counting Go No-Go Omissions.

Overall, results indicate that ADHD children generally demonstrated poorer executive functioning across behavior and performance measures of EF, thus supporting
hypothesis 1; however, scores on EF performance measures were more variable. In addition, despite differences between ADHD-I and ADHD-C/HI on the BRIEF Working Memory and Inhibit scales, the two ADHD groups tended to perform similarly on the EF performance measures reflective of these constructs. Thus Hypotheses 2 and 3 were not supported for the performance-based measures as they were for the BRIEF rating scales.

Multitrait Multimethod Analyses

Testing Hypotheses 4 and 5 involved Multitrait Multimethod analyses designed to test the following predictions:

- That ratings on the BRIEF would be more strongly correlated with performance-based measures of EF overall than with measures of intellectual functioning (Hypothesis 4)
- That performance and behavior checklist measures of similar EF constructs would correlate more strongly with each other than with measures of theoretically different aspects of EF assessed by the same methods (Hypothesis 5).

In order to test these hypotheses, two steps were required: First, it was necessary to obtain correlations of the BRIEF scales with FSIQ and the 10 performance measures of EF (i.e., the 4 D-KEFS measures, Spatial Span, Conflicting Motor Response, Lateral Gaze, CANTAB Self-Ordered Pointing, and the two Go No-Go tasks). Second, it was necessary to test the magnitude of the correlations in order to evaluate convergent and divergent validity of the measures.

Correlations of BRIEF, EF performance measures, and FSIQ: Recall that correlations of IQ with all EF behavior rating and EF performance measures were presented in Appendix B. Correlations between the BRIEF MCI, Working Memory, and BRIEF Inhibit measures were highly correlated with FSIQ. FSIQ was also significantly correlated with 5 of the 10 EF performance measures (WISC Spatial Span Backward, D-
KEFS Trail Making Test, D-KEFS Category Switching, D-KEFS Color-word Interference Test Inhibition, and Conflicting Motor Response).

Partial correlations (controlling for age and IQ) between the BRIEF and the performance-based EF measures are listed in Table 7. An alpha level of 0.01 and a medium effect size was used to determine significance; however, for descriptive purposes, effect sizes with $p$-values less than 0.05 are also indicated with asterisks in Table 7.

### Table 7. Correlations of BRIEF and EF performance measures

<table>
<thead>
<tr>
<th></th>
<th>Inhibit</th>
<th>BRI</th>
<th>WM</th>
<th>MCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$d$</td>
<td>$r$</td>
<td>$d$</td>
</tr>
<tr>
<td>Trails 4</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.12</td>
<td>0.24</td>
</tr>
<tr>
<td>CF Switching</td>
<td>-0.08</td>
<td>0.16</td>
<td>-0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>CWI Inhibition</td>
<td>0.06</td>
<td>0.12</td>
<td>0.27</td>
<td>0.56</td>
</tr>
<tr>
<td>Tower</td>
<td>-0.14</td>
<td>0.29</td>
<td>-0.19</td>
<td>0.39</td>
</tr>
<tr>
<td>Span Backwards</td>
<td>-0.07</td>
<td>0.13</td>
<td>-0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>Conflicting Motor.</td>
<td>-0.29</td>
<td>0.6</td>
<td>-0.29</td>
<td>0.6</td>
</tr>
<tr>
<td>Lateral Gaze</td>
<td>-0.46</td>
<td>***1.04</td>
<td>-0.6</td>
<td>***1.48</td>
</tr>
<tr>
<td>SOP Between Errors</td>
<td>0.14</td>
<td>0.27</td>
<td>0.21</td>
<td>0.43</td>
</tr>
<tr>
<td>Simple GNG</td>
<td>0.25</td>
<td>0.51</td>
<td>0.23</td>
<td>0.47</td>
</tr>
<tr>
<td>Counting GNG</td>
<td>-0.29</td>
<td>0.61</td>
<td>-0.27</td>
<td>0.55</td>
</tr>
</tbody>
</table>

**Average r value:** 0.18

**Average d value:** 0.377

Note. *$p < .05$; **$p < .01$; ***$p < .001$; BRIEF = Behavior Rating Inventory of Executive Function; Inh = Inhibition Scale; BRI = Behavioral Regulation Index; WM = Working Memory Scale; MCI = Metacognition Index; Trails 4 = D-KEFS Trails Condition 4 Total Time; CF Switching = D-KEFS Category Switching Total; CWI Inhibition = D-KEFS Color Word Interference Inhibition; Tower Achievement = D-KEFS Tower Total Achievement; Span Backwards = WISC Spatial Span Backwards; Conflicting Motor. = Conflicting Motor Response; Lateral Gaze = Lateral Gaze Total Time; SOP Between Errors = Self-Ordered Pointing Between Errors; Simple Go No-Go = Simple Go No-Go Commissions; Counting Go No-Go = Counting Go No-Go Omissions.

Of the EF measures, the BRIEF Working Memory was significantly correlated with Conflicting Motor Response, Lateral Gaze, and Counting Go No-Go (1 of 5 working memory measures, and 2 of 5 inhibition measures), while the BRIEF MCI was correlated
with Conflicting Motor Response, Lateral Gaze, SOP Between Errors, Simple Go No-Go, and Counting Go No-Go (2 of 5 WM measures, and 3 of 5 IC measures). The BRIEF Inhibit scale was not correlated with any measures at the .01 significance level, and the BRIEF BRI was correlated only with Lateral gaze. However, several correlations with the BRIEF Inhibit Scale and BRI, as well as the Working Memory scale, approached significance (e.g. below the .05 level of significance).

**Evaluation of convergent and divergent validity**

Z-tests were conducted to analyze whether the relevant correlations obtained from the analyses described above were significantly similar or different from one another to support convergent and divergent validity of the EF measures included in this investigation. All correlations were transformed to z-scores using Fisher’s r to z transformations and z-tests were performed on the resultant z-scores for the BRIEF with each combination of IQ and 10 EF performance measures.

Examining BRIEF correlations with IQ and EF performance measures: In order to test Hypothesis 4, mean correlations were calculated for correlations of the 4 BRIEF measures (Inhibit, WM, BRI, and MCI) with IQ and with each of the 10 EF performance measures. These correlations were compared using z-tests to determine whether these measures from the BRIEF were correlated more strongly with EF performance measures than with IQ. Z-tests and p-values for the IQ comparisons are presented in Table 8.

While the correlation between the BRIEF and IQ was larger than all but one of the correlations between the BRIEF and the performance measures (Lateral Gaze), z-test comparisons indicated that the magnitudes of these correlations were not significantly different overall at the conservative p-value of 0.01. However, 4 of the 10 comparisons
were close to significance, suggesting that the correlations between the BRIEF and IQ were slightly stronger than correlations between the BRIEF and some of the EF measures.

Examination of the IQ correlations with all EF rating scale and performance measures, listed in Appendix B, and examination of the correlations between the BRIEF and EF performance measures, listed in Table 8, indicated that IQ was similarly correlated with all EF measures while correlations between the BRIEF and EF performance measures were more variable. In particular, the BRIEF was weakly more correlated with the D-KEFS measures than the unstandardized neurobehavioral measures.

Table 8. Comparisons of EF correlations with IQ

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>r vs. BRIEF x FSIQ</th>
<th>z-score</th>
<th>2 tailed p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIEF x Trails 4</td>
<td>0.04</td>
<td>0.39</td>
<td>2.12</td>
<td>0.03</td>
</tr>
<tr>
<td>BRIEF x CF Switching</td>
<td>0.06</td>
<td>0.39</td>
<td>2.00</td>
<td>0.04</td>
</tr>
<tr>
<td>BRIEF x CWI Inhibition</td>
<td>0.19</td>
<td>0.39</td>
<td>1.25</td>
<td>0.21</td>
</tr>
<tr>
<td>BRIEF x Tower</td>
<td>0.13</td>
<td>0.39</td>
<td>1.60</td>
<td>0.06</td>
</tr>
<tr>
<td>BRIEF x Span Backwards</td>
<td>0.13</td>
<td>0.39</td>
<td>1.62</td>
<td>0.05</td>
</tr>
<tr>
<td>BRIEF x Contralateral Motor Resp.</td>
<td>0.31</td>
<td>0.39</td>
<td>0.75</td>
<td>0.45</td>
</tr>
<tr>
<td>BRIEF x Lateral Gaze</td>
<td>0.40</td>
<td>0.39</td>
<td>0.27</td>
<td>0.79</td>
</tr>
<tr>
<td>BRIEF x SOP Between Errors</td>
<td>0.24</td>
<td>0.39</td>
<td>0.94</td>
<td>0.34</td>
</tr>
<tr>
<td>BRIEF x Simple GNG</td>
<td>0.31</td>
<td>0.39</td>
<td>0.51</td>
<td>0.61</td>
</tr>
<tr>
<td>BRIEF x Counting GNG</td>
<td>0.34</td>
<td>0.39</td>
<td>0.30</td>
<td>0.76</td>
</tr>
<tr>
<td>Average z score</td>
<td></td>
<td></td>
<td>1.14</td>
<td>0.33</td>
</tr>
<tr>
<td>Average z score of comparisons with DKEFS measures</td>
<td></td>
<td></td>
<td>1.74</td>
<td>0.08</td>
</tr>
<tr>
<td>Average z score of comparisons with other neurobehavioral measures</td>
<td></td>
<td></td>
<td>0.55</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Note. BRIEF = Behavior Rating Inventory of Executive Function; Inh = Inhibition Scale; BRI = Behavioral Regulation Index; WM = Working Memory Scale; MCI = Metacognition Index; Trails 4 = D-KEFS Trails Condition 4 Total Time; CF Switching = D-KEFS Category Switching Total; CWI Inhibition = D-KEFS Color Word Interference Inhibition; Tower Achievement = D-KEFS Tower Total Achievement; Span Backwards = WISC Spatial Span Backwards; Conflicting Motor = Conflicting Motor Response; Lateral Gaze = Lateral Gaze Total Time; SOP Between Errors = Self-Ordered Pointing Between Errors; Simple Go No-Go = Simple Go No-Go Commissions; Counting Go No-Go = Counting Go No-Go Omissions.
Consequently, Hypothesis 4 was not supported in two ways a) the magnitude of correlations between IQ and the EF rating scale and performance measures were not significantly different from one another, and b) correlations between the BRIEF and EF performance measures were more variable, and in many cases, smaller than correlations between IQ and all EF rating scale and performance measures. Examining BRIEF Working Memory and Inhibit scale correlations with EF performance measures: In order to evaluate whether WM measures better correlated with one another than with measures of IC, and conversely, whether measures of IC better correlated with one another than with WM measures (Hypothesis 5), z-tests were conducted to examine correlations among WM and IC measures to further evaluate the convergent and divergent validity of the EF measures included in this investigation. Correlations of the BRIEF Working Memory Scale with the 10 EF performance measures were considered along with correlations of the BRIEF Inhibit Scale with the 10 EF performance measures. Significance was evaluated through z-tests to determine whether the BRIEF Working Memory Scale and performance measures of WM correlated more strongly with one another than with measures of IC, and conversely whether the BRIEF Inhibit Scale and performance measures of IC correlated more strongly with one another than with measures of other EF constructs assessed by similar methods.

Z-tests and p-values for the WM and IC comparisons are presented in Table 9. Only two of the EF comparisons yielded significant differences between the sets of correlations, Lateral Gaze (Inhibitory control) and the CANTAB Self-Ordered Pointing tests (Working Memory).
Thus, Hypothesis 5 was not supported as measures of WM and IC correlated similarly to one another. In addition, examination of the correlations between BRIEF and EF performance measures in Table 7 indicate that the Working Memory scale was more highly correlated with both Lateral Gaze ($r = -0.48$ vs. $r = -0.26$) and Self-Ordered Pointing ($r = 0.3$ vs. $r = 0.07$). While Self-Ordered Pointing was expected to correlate more strongly with the BRIEF Working Memory scale, the stronger correlation of Lateral Gaze with the BRIEF Working Memory scale than the Inhibit scale was somewhat surprising. A possible explanation, to be explored in the discussion section, was that the performance and rating scale measures of IC likely assessed different aspects of behavioral inhibition.

**Table 9. Comparisons of Working Memory and Inhibitory Control measures**

<table>
<thead>
<tr>
<th>Measure</th>
<th>$r$</th>
<th>Measure</th>
<th>$r$</th>
<th>$z$-score</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WM x Trails 4</td>
<td>0.11</td>
<td>Inh x Trails 4</td>
<td>0.07</td>
<td>0.22</td>
<td>0.41</td>
</tr>
<tr>
<td>WM x CF Switching</td>
<td>0.09</td>
<td>Inh x CF Switching</td>
<td>0.05</td>
<td>0.23</td>
<td>0.41</td>
</tr>
<tr>
<td>WM x CWI Inhibition</td>
<td>0.20</td>
<td>Inh x CWI Inhibition</td>
<td>0.15</td>
<td>0.27</td>
<td>0.39</td>
</tr>
<tr>
<td>WM x Tower</td>
<td>0.15</td>
<td>Inh x Tower Achievement</td>
<td>0.16</td>
<td>0.03</td>
<td>0.49</td>
</tr>
<tr>
<td>WM x Span Backwards</td>
<td>0.17</td>
<td>Inh x Span Backwards</td>
<td>0.08</td>
<td>0.51</td>
<td>0.30</td>
</tr>
<tr>
<td>WM x Conflicting Motor</td>
<td>0.30</td>
<td>Inh x Conflicting Motor</td>
<td>0.27</td>
<td>0.14</td>
<td>0.44</td>
</tr>
<tr>
<td>WM x Lateral Gaze</td>
<td>0.48</td>
<td>Inh x Lateral Gaze</td>
<td>0.26</td>
<td>1.44</td>
<td>0.08</td>
</tr>
<tr>
<td>WM x SOP Between Errors</td>
<td>0.30</td>
<td>Inh x SOP Between Errors</td>
<td>0.07</td>
<td>1.31</td>
<td>0.09</td>
</tr>
<tr>
<td>WM x Simple GNG</td>
<td>0.28</td>
<td>Inh x Simple GNG</td>
<td>0.24</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>WM x Counting GNG</td>
<td>0.41</td>
<td>Inh x Counting GNG</td>
<td>0.21</td>
<td>1.12</td>
<td>0.13</td>
</tr>
<tr>
<td>Average $z$ score</td>
<td>0.55</td>
<td></td>
<td>0.19</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>Average $z$ score of comparisons with DKEFS measures</td>
<td>0.80</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. BRIEF = Behavior Rating Inventory of Executive Function; Inh = Inhibition Scale; WM = Working Memory Scale; Trails 4 = D-KEFS Trails Condition 4 Total Time; CF Switching = D-KEFS Category Switching Total; CWI Inhibition = D-KEFS Color Word Interference Inhibition; Tower Achievement = D-KEFS Tower Total Achievement; Span Backwards = WISC Spatial Span Backwards; Conflicting Motor = Conflicting Motor Response; Lateral Gaze = Lateral Gaze Total Time; SOP Between Errors = Self-Ordered Pointing Between Errors; Simple Go No-Go = Simple Go No-Go Commissions; Counting Go No-Go = Counting Go No-Go Omissions.
Multitrait-Multimethod Matrix: Correlations between all rating scales and performance measures are provided in Appendix C. Figure 5 (below) displays the results of these correlations in a Multitrait-Multimethod Matrix (MMM) for the three measures that provided reasonable assessment of WM and IC within a single method (e.g., the BRIEF, DKEFS, and Go No-Go tasks). Correlations along the top diagonal are the reliability coefficients and should be the highest in the MMM as a construct should more highly correlated with itself than with anything else. Note, however, that while the correlations listed for the BRIEF and GNG are perfectly correlated at 1.0 because they are correlations between identical scales, the correlations listed for the D-KEFS are

<table>
<thead>
<tr>
<th>Traits</th>
<th>BRIEF</th>
<th>DKEFS</th>
<th>Go No-Go</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( WM_1 )</td>
<td>( I_1 )</td>
<td>( WM_2 )</td>
</tr>
<tr>
<td>BRIEF</td>
<td>1.0</td>
<td>0.64</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>( I_1 )</td>
<td>1.0</td>
<td>0.13</td>
</tr>
<tr>
<td>DKEFS</td>
<td>( WM_2 )</td>
<td>( I_2 )</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>( I_3 )</td>
<td>0.33</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Figure 5. Multitrait-Multimethod Matrix of Working Memory and Inhibitory Control
notably below 1.0 as these are the correlations of the D-KEFS Category Switching with the D-KEFS Tower for the working memory construct and of the D-KEFS Trails and the D-KEFS Color-Word Interference for the inhibitory control construct. Still, these correlations should be higher than the remaining correlations if they were measures of similar constructs assessed by similar methods (homotrait-homomethod) as hypothesized. In contrast, the lowest correlations in the matrix should be those in the triangles (heterotrait-hetermethod) because they are correlations between different constructs assessed by different methods. The remaining correlations of heterotrait-homomethod, displayed between the triangles, should fall somewhere between homotrait-homomethod correlations and heterotrait-hetermethod correlations because they are correlations of different constructs measured by similar methods. Thus, support for construct validity is supported by a general rank ordering of the correlation coefficients with reliability > convergent validity > divergent validity, heterotrait-monomethod > divergent validity, heterotrait-hetermethod. It is important to remember that that interpretation of the MMM requires some interpretation and judgment as it rare that all precise rank ordering is possible.

Initial examination of the correlations displayed in the MMM confirmed the above analyses that convergent and divergent validity of the working memory and inhibitory control constructs were not supported by the measures listed here. Overall, method variance appeared to be stronger than construct validity. However, exploration of all the correlations provides some limited support for convergent and discriminant validity. To summarize correlations provided in Appendix C, average correlations among WM and IC measures have been listed below in Table 10 and Table 11.
A review of the correlations indicated that the BRIEF Working Memory scale generally correlated more highly with performance measures than the BRIEF Inhibit Scale with performance measures, whereas correlations within the IC performance measures were generally higher than correlations within the WM performance measures. Possible reasons for this pattern of results to be explained in the Discussion section include differences between the assessment methods and in the aspects of WM and IC assessed.

Table 10. Average Correlations for Inhibitory Control Measures

<table>
<thead>
<tr>
<th>Correlations of IC Performance Measures</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-KEFS (Trails &amp; Color Word)</td>
<td>0.24</td>
</tr>
<tr>
<td>Neurobehavioral (Conflicting, Lateral, Simple)</td>
<td>0.28</td>
</tr>
<tr>
<td>All IC Performance Measures</td>
<td>0.25</td>
</tr>
<tr>
<td>Correlations of BRIEF Inhibit Scale with</td>
<td></td>
</tr>
<tr>
<td>BRI</td>
<td>0.94</td>
</tr>
<tr>
<td>D-KEFS IC Performance measures</td>
<td>0.04</td>
</tr>
<tr>
<td>Neurobehavioral Performance measures</td>
<td>0.33</td>
</tr>
<tr>
<td>All IC Performance measures</td>
<td>0.22</td>
</tr>
<tr>
<td>All Performance measures (both WM and IC)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 11. Average Correlations for Working Memory measures

<table>
<thead>
<tr>
<th>Correlations of WM Performance Measures</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-KEFS (Category &amp; Tower)</td>
<td>0.04</td>
</tr>
<tr>
<td>Neurobehavioral (Spatial, Self-ordered Pointing, Counting)</td>
<td>0.18</td>
</tr>
<tr>
<td>All WM Performance measures</td>
<td>0.11</td>
</tr>
<tr>
<td>Correlations of BRIEF Working Memory Scale with</td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>0.95</td>
</tr>
<tr>
<td>D-KEFS WM Performance measures</td>
<td>0.17</td>
</tr>
<tr>
<td>Neurobehavioral Performance measures</td>
<td>0.28</td>
</tr>
<tr>
<td>All WM Performance measures</td>
<td>0.23</td>
</tr>
<tr>
<td>All Performance Measures (both WM and IC)</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Additional Analyses

Although the above analyses did not provide clear evidence of convergent and divergent validity for the WM and IC constructs across the rating scale and performance measures, rating scale measures of these two constructs did provide discriminant validity between the ADHD subtypes. Table 12 lists the correlations of the four BRIEF measures with the two CPRS scales. All correlations were highly significant \((p < 0.0001)\). A cursory inspection of the correlations suggests that the BRIEF BRI and Inhibit scale were more highly correlated with the DSM criteria of hyperactivity-impulsivity from the CPRS Scale M, than with the DSM criteria of inattention from the CPRS Scale L. Conversely, the BRIEF MCI and Working Memory scale appeared more highly correlated with the DSM criteria of inattention from the CPRS Scale L, than with the DSM criteria of hyperactivity-impulsivity from the CPRS Scale M.

Table 12. Correlations between the BRIEF and CPRS

<table>
<thead>
<tr>
<th>BRIEF Inhibit</th>
<th>BRIEF BRI</th>
<th>BRIEF WM</th>
<th>BRIEF MCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>d</td>
<td>r</td>
<td>d</td>
</tr>
<tr>
<td>CPRS-I</td>
<td>0.58</td>
<td>***1.44</td>
<td>0.59</td>
</tr>
<tr>
<td>CPRS-HI</td>
<td>0.78</td>
<td>***2.46</td>
<td>0.66</td>
</tr>
<tr>
<td>Average r value</td>
<td>0.68</td>
<td></td>
<td>0.63</td>
</tr>
<tr>
<td>Average d value</td>
<td>1.95</td>
<td></td>
<td>1.60</td>
</tr>
</tbody>
</table>

NOTE: BRIEF = Behavior Rating Inventory of Executive Function; Inh = Inhibition Scale; BRI = Behavioral Regulation Index; WM = Working Memory Scale; MCI = Metacognition Index; CPRS – I = Conners Parent Rating Scale DSM Inattentive (Scale L); CPRS HI = Conners Parent Rating Scale DSM Hyperactive-Impulsive (Scale M).

To examine whether these correlations were significantly different from one another, all correlations were transformed to \(z\)-scores using Fisher’s \(r\) to \(z\) transformations and \(z\)-tests were performed on the resultant \(z\)-scores. Table 13 lists the \(z\)-tests for correlations between the four BRIEF measures with the two CPRS scales. Results
indicate the BRIEF Inhibit scale was more strongly correlated with CPRS Hyperactivity-Impulsivity symptoms \((r = .78)\) than with CPRS Inattention symptoms \((r = .58)\), while the Working Memory Scale was more strongly correlated with CPRS Inattention symptoms \((r = .91)\) than with CPRS Hyperactivity/Impulsivity symptoms \((r = .66)\). In addition, the BRIEF MCI was more strongly correlated with CPRS Inattention symptoms \((r = .91)\) than with the CPRS Hyperactivity/Impulsivity symptoms \((r = .71)\); however, the BRIEF BRI was not more strongly correlated with CPRS Hyperactivity/Impulsivity symptoms \((r = .66)\) than with CPRS Inattention symptoms \((r = .59)\). These results support the theoretical impression that deficits in WM and IC are differentially associated with the behavioral symptom clusters of ADHD. Further, since these data include children with and without ADHD, these results suggest that aspects of WM are more strongly associated with attentional control processes whereas IC is more closely associated with aspects of behavioral regulation.

Table 13. Comparisons of the BRIEF and CPRS

<table>
<thead>
<tr>
<th></th>
<th>(r)</th>
<th>vs.</th>
<th>(r)</th>
<th>(z) score</th>
<th>2 tailed (p) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inh x HI</td>
<td>0.78</td>
<td></td>
<td>0.58</td>
<td>2.18</td>
<td>0.01</td>
</tr>
<tr>
<td>BRI x HI</td>
<td>0.66</td>
<td></td>
<td>0.59</td>
<td>0.66</td>
<td>0.26</td>
</tr>
<tr>
<td>WMI x I</td>
<td>0.91</td>
<td></td>
<td>0.66</td>
<td>4.19</td>
<td>0.00</td>
</tr>
<tr>
<td>MCI x I</td>
<td>0.91</td>
<td></td>
<td>0.71</td>
<td>3.65</td>
<td>0.00</td>
</tr>
</tbody>
</table>

NOTE: BRIEF = Behavior Rating Inventory of Executive Function; Inh = Inhibition Scale; BRI = Behavioral Regulation Index; WM = Working Memory Scale; MCI = Metacognition Index; CPRS = Conners Parent Rating Scale, I = CPRS DSM Inattentive (Scale L); HI = CPRS DSM Hyperactive-Impulsive (Scale M).
CHAPTER 4
DISCUSSION

Executive function (EF) has been become a popular area of study in pediatric neuropsychology. The parallels between the symptom characteristics of many neurodevelopmental disorders and delayed maturation of the frontal lobes through childhood and into early adulthood has lead many investigators to emphasize the importance of executive dysfunction (EdF) in the developmental outcome and the recovery from cerebral disorders affecting the developing brain. However, our understanding of the development of EF skills in children and the potential for disruption of these abilities due to brain pathology is relatively limited, and rather late in emerging considering the long history of investigations of frontal lobe pathology in adults. While it is certainly plausible that development of EF is an important part of a child’s cognitive, behavioral, and social-emotional development, the impact of executive dysfunction during childhood remains uncertain and much debated, particularly what constitutes impaired EF within the context of a developing brain and what are the critical periods for acquisition of EF skills during development. Consequently, appropriate and ecologically valid assessment of EF is critical to advancing our understanding of EF in children.

This study examined the ecological validity of EF measures in children with and without Attention Deficit Hyperactivity Disorder (ADHD). In particular, the divergent and convergent validity of working memory (WM) and inhibitory control (IC) were evaluated through the use of behavior rating scales and several performance measures of executive function. The predictive validity of EF measures in differentiating between
children with and without ADHD was explored with particular emphasis on the ability of WM and IC measures to differentiate between children with ADHD. The constructs of WM and IC were chosen because of their primary role in models of EF (Fuster, 1999; Cohen and Servan-Schreiber, 1992) and in theories about the core deficits and underlying neuropathology of ADHD (Barkley, 1997; Pennington et al., 1997).

Two sets of hypotheses were proposed. First, it was predicted that EF measures would differentiate between children with and without ADHD, and further, that measures of WM and IC would differentiate between subtypes of ADHD. Multivariate analyses of covariance, controlling for age and IQ, were used to investigate group differences on these measures. The second set of hypotheses involved investigation of the construct validity of WM and IC. It was hypothesized that measures of working memory would correlate more strongly with one another than with measures of inhibitory control or intellectual functioning, and conversely, that measures of inhibitory control would correlate more strongly with one another than with measures of working memory or intellectual functioning.

In brief, results of analyses of covariance revealed group differences between children with and without ADHD on rating scales measures of EF, but inconsistent group differences on performance measures of EF. Behavior ratings of WM and IC also better differentiated between ADHD subtypes than performance measures. In addition, correlations were low to moderate between EF rating scales and performance measures. Discussion of the interpretation and implications of these findings below will consider several points, including: the known variability across studies in identifying group differences in EF performance, difficulties associated with assessment of EF through
performance measures and whether rating scale and performance measures assess similar aspects of EF, consideration of whether WM and IC are independent constructs, potential differences in the range of scores resulting in a restriction of range, and the confounding impact of the IQ and age effects in this sample.

Findings from this study provide support for the use of the Behavior Rating Inventory of Executive Function, (BRIEF; Gioia et al., 2000) in identifying the diagnostic groups of ADHD-Inattentive type (ADHD-I) and Combined type (ADHD-C/HI) as initially demonstrated by Gioia et al. (2002) and Mahone et al. (2002). Children with ADHD-I and ADHD-C/HI were both rated higher on the BRIEF Working Memory scale and Metacognition Index (MCI) but only children with ADHD-C/HI were rated higher on the Inhibit scale and Behavior Regulation Index (BRI) of the BRIEF. In addition, the Working Memory scale and MCI were more strongly correlated with symptoms of inattention than with symptoms of hyperactivity and impulsivity, whereas the Inhibit scale, and to a moderate degree, the BRI, was more strongly correlated with symptoms of hyperactivity and impulsivity.

These findings support the utility of applying a neuropsychological model of executive function to the behavioral symptomatology of ADHD. Deficits in WM were associated with symptoms of inattention and predictive of both subtypes of ADHD whereas only IC was a strong factor in the prediction of the Combined type and was more strongly associated with symptoms of hyperactivity and impulsivity. While this study only examined the constructs of WM and IC, it is worth noting that Gioia et al. (2002), using the commercially available version of BRIEF with a two factor structure, found a similar pattern of association for the other scales of the MCI (initiation, planning,
organization of materials, and monitor) and other measures of inattention, including the Behavior Assessment System for Children (Reynolds & Kamphaus, 1994), Child Behavior Checklist (Achenbach, 1991), and the ADHD Rating Scale IV (DuPaul, Power, Anastopoulos & Reid, 1998). Hyperactive-impulsive symptoms as well as measures of aggression were also more strongly correlated with the other two scales of the BRI (Shift and Emotional control), although these two scales correlated more weakly than the Inhibit scale. Gioia et al. (2002) also noted that the Working Memory and Inhibit scales correlated moderately with scales reflecting social difficulties, consistent with the observation that children who have attention and, in particular, impulse control problems, also have social difficulties. In addition, there were lower correlations on scales assessing behavioral and emotional difficulties (e.g., anxiety, depression somatic complaints, conduct problems, atypicality, etc.) that should not be related to inattention, impulsivity, and hyperactivity. The pattern of low correlations with scales measuring somatic complaints, anxiety, and depression provided evidence of divergent validity for the Working Memory and Inhibit scales, while the pattern of high correlations with scales measuring inattention, hyperactivity, and impulsivity provided further support that deficits in the domains of executive function are strongly related to the clinical diagnosis of Attention-Deficit/Hyperactivity Disorder, as suggested by many (Barkley, 1997, Denckla, 1996).

In contrast to the strong findings from the BRIEF, the performance measures of EF did not differentiate as consistently between the ADHD diagnostic groups and controls. In addition, while it was hoped that use of a well-standardized measure such as the D-KEFS would provide more sensitivity in detecting EF impairments in ADHD children,
results suggested that the unstandardized neurobehavioral measures were better at differentiating between children with and without ADHD, particularly those that assessed IC. Further, although there were differences between the groups on Conflicting Motor Response, Lateral Gaze, D-KEFS Color Word Interference, D-KEFS Tower, Self-Ordered Pointing, and Counting Go No-Go Omissions, the overall performance of ADHD children was within normal limits as indicated by the normative scores for the D-KEFS in the average range and scores on all measures within in one standard deviation of the mean, suggesting that while the performance of ADHD children was significantly different from this control group from a statistical perspective, performance on these measures was not necessarily impaired, but only poorer.

While executive dysfunction in children with ADHD has been a reliable finding, the exact measures that differentiate controls from children with ADHD vary considerably across studies (Pennington and Ozonoff, 1996). In addition, while findings from performance-based measures may statically differentiate between groups, they do not by themselves reliably differentiate between subtypes of ADHD (Dykman & Ackerman, 1993). Baron (2003) notes that it is common for children to exhibit normal performance on EF measures in a clinic setting while demonstrating executive dysfunction in the real world. These findings suggest that lower scores on performance measures of EF tests may occur due to ADHD, but may lack specificity. In addition, normal scores are not indicative of absence of ADHD and may be uninterpretable.

To provide a foundation for evaluating construct validity across the EF measures, the Multitrait-Multimethod methodology proposed by Campbell and Fiske (1959) was employed to facilitate examination of convergent and divergent validity in conjunction
with investigation of variance associated with the method of administration (method variance). The arrangement of correlations between constructs and methods in a Multitrait-Multimethod Matrix (MMM) allows detection of variance unique to each construct and variance unique to each method of assessment. Consistent with previous findings (Stuss & Benson, 1984, Walsh, 1985), the relationship between behavioral rating and performance measures of EF was inconsistent in this study. Correlations across the two methods of assessment suggested stronger variance attributed to method than construct. Moreover, there was not a consistent pattern of correlations between measures of theoretically similar constructs, indicating that the construct validity of WM and IC was not supported across the behavior rating and performance measures. Correlations between the BRIEF Working Memory scale and all performance measures were generally stronger than correlations between the BRIEF Inhibit Scale with all performance measures, whereas correlations among the IC performance measures only were generally higher than correlations among the WM performance measures.

Campbell and Fiske (1959) considered several alternative explanations in cases where support for construct validity appears to be lacking. The first possibility is that none of the methods used were adequate for assessment of EF; however, given the research cited for the behavior rating scales and the performance measures employed in this study (presented in the measures section in chapter 2), it seemed reasonable to conclude that these measures were appropriate for assessment of EF, at least at the global level of executive functioning.

A second alternative is the possibility that one of the two methods (rating scale or performance measures) inadequately assessed the specific constructs of WM and IC, thus
limiting the ability of these methods to provide corresponding measurement of working
memory and inhibitory control. This alternative can be considered a more likely
possibility due to differences between the two assessment methods in the setting and the
window of time for evaluation of the behaviors, as well as differences between the
methods in the degree of specificity for measuring WM and IC. Sufficient research has
supported the use of the behavior rating and performance measures employed in this
study to assess EF; however, the sampling of behaviors afforded by these two methods
assessed behavioral functioning under different levels of structure. Parent ratings on the
BRIEF were based upon observations of the child’s behavior across time and in less
predictable and less structured settings than the clinic. It has been suggested by many that
the structure provided in a clinical assessment decreases demands on EF processes
(Bernstein & Waber, 1990; Stuss & Benson, 1986; Kaplan, 1988), thus limiting the
ability of performance measures to detect weaknesses in EF. In the assessment
environment, the examiner often provides support and encouragement (Sbordone, 2000),
and plans, initiates, and paces activities (Anderson, 2002). The child may be able to more
effectively self-regulate in the testing environment while working one-on-one with an
examiner and under clear rules. Although the degree of structure was not objectively
assessed in this study, it is likely that the examiner provided a good deal of structure as
children were encouraged to continue working steadily through the 2 ½ hour assessment.
Care was taken to switch between computer and paper-pencil tasks in order to provide a
number of opportunities to move around the room and to vary the level of activity,
particularly in the latter half of the session when children switched more frequently
between shorter tasks. In addition, the assessment may not have been sufficient in length
or stressful enough to elicit executive dysfunction, as children were typically actively engaged and motivated throughout assessment.

Behavioral rating scales and performance measures may also tap fundamentally different aspects of EF by virtue of the assessment methods particular to each measure. Although the utility of behavior rating scales in describing real-world functioning has been well established (Burgess, 1997), behavioral testing of EF in the clinic or laboratory has been demonstrated to be inconsistent with observer ratings of the behavior examined (Eslinger & Damasio, 1985; Levine et al., 1998; Mesulam, 1986; Stuss & Buckle, 1992). It is possible that behavior rating scales of EF are better suited for assessment of social-emotional consequences of executive dysfunction while performance measures are better for detecting deficits in specific motor output that are assumed to reflect internal executive control processes. While both methods assess aspects of behavioral regulation, observer ratings of EF focus more on the informant’s impression of socio-emotional consequences of the child’s executive dysfunction and the impact of the child’s behavior on others and their environment, whereas performance measures quantify the specific behavioral output of a presumed cognitive process without regard for the socio-emotional aspects of EF. In practice, an examiner could observe these socio-emotional behaviors during performance testing and describe them in clinically, but they are not typically measured in any objective manner. Consequently, it is possible for children to perform within normal limits but with varying degrees of discomfort and difficulty observed during their performance that does not impact the final score.

An examination of the correlations between the BRIEF and the performance measures revealed an additional insight into the possibility that the BRIEF assessed
slightly different aspects of EF than performance measures. Overall, the Working Memory scale was more highly correlated across all types of measures ($r = .25$) (behavioral ratings and performance measures) than the BRIEF Inhibit scale ($r = .18$) (refer to Tables 10 & 11). Exploration of the individual items from these scales suggested that the Working Memory scale may have more adequately captured cognitive aspects of inattention and working memory difficulties (e.g., Has a short attention span, Has trouble concentrating on chores, schoolwork, etc.; Is easily distracted by noises, activity, sights, etc.; When sent to get something, forgets what he/she is supposed to get) than the Inhibit scale, which contained items that primarily assess aspects of social disinhibition (Acts wilder or sillier than others in groups; Interrupts others; Gets out of seat at the wrong times; Gets out of control more than friends; Acts too wild or "out of control"; Becomes too silly). Performance measures of working memory commonly assess the behaviors included among the items of BRIEF Working Memory scale (e.g., memory span) whereas the performance measures of inhibitory control tend to assess aspects of motor control rather than social-emotional behaviors. In addition, it should also be considered that performance measures may tap cognitive aspects of EF that are more closely linked to dorsolateral regions of the prefrontal cortex, whereas rating scales better capture behavioral aspects of EF that are more dependent upon orbital and ventral-medial areas (Anderson, 2002). Combining these two assertions, it reasonable to see how measures of working memory correlated more highly across rating scales and performance measures as both appeared to adequately assess cognitive aspects of WM whereas only the performance measures may have adequately assessed cognitive aspects of IC while the rating scales assessed social-emotional aspects of IC. It would be
interesting to investigate this possibility further in pediatric samples where isolation of these prefrontal areas is possible.

There may also be differences in the ability of behavior rating and performance measures to detect discrete EF processes (Baron, 2003). Given that EF measures are “multi-determined” assessments dependent upon the integrity of lower-order cognitive skills (Denckla, 1996) that are inevitably integrated at higher levels of cognition for behavioral output (Burgess, 1997), several investigators have concluded that laboratory and clinical measures of EF lack sensitivity to assess the multidimensional nature of EF (Anderson, 2002). Pennington et al. (1996) stated that while EF measures assess a wide range of meaningful variance, it is difficult to know which cognitive processes account for a given score. In contrast, rating scale measures may contain items that more precisely define and assess behaviors associated with executive dysfunction, thus isolating EF processes at a descriptive level that are difficult to functionally isolate through performance measures (Gioia et al., 2001).

Anderson (2002) suggested that measures focused on isolating strategic components of EF might better detect specific deficits in EF and demonstrate ecological validity with real-world functioning. Her assertion may be supported by results from this study as the medium sized correlations between the BRIEF and the unstandardized neurobehavioral measures were larger than correlations between the BRIEF and the D-KEFS, indicating that is possible to obtain convergence across the behavioral and performance measures when sensitivity and specificity are adequate. The unstandardized neurobehavioral measures could be argued to be less complex, and by extension, more precise in isolating particular aspects of EF, given the relatively fewer operational
requirements and rules of these tasks compared to the D-KEFS measures. For example, Lateral Gaze and Conflicting Motor Response both could be considered relatively purer measures of inhibitory control with low cognitive demands, compared to the D-KEFS measures that required utilization of multiple operational rules, monitoring of response over time (e.g., no repetitions, planning a sequence of moves) and inhibition of incorrect answers. Thus, it is possible that the BRIEF correlated more strongly with the unstandardized neurobehavioral measures due to greater specificity of the constructs assessed by the measures.

It is also worth noting here that one of the perceived strengths of the D-KEFS measure, in addition to being nationally normed using census data, was the use of “process” scores that provide quantitative measures of more qualitative aspects of EF performance. These process scores could serve as the “strategic” parameters that Anderson (2002) suggested might provide better measures of EF. Process scores have theoretical support in adults for evaluating components of EF and separating performance on subdomains of cognitive functioning processes from EF abilities (Delis et al., 2003). Although not a focus of this study, these scores (e.g. set-loss errors, repetition errors, subtest contrast scores, etc.) were analyzed to explore the utility of these scores in differentiating between ADHD and control groups. Although testing observations seemed to indicate that children with ADHD made more errors, children of all groups made few mistakes and NONE of the process scores yielded significant differences between ADHD and control groups. Future investigations from this data will evaluate whether use of these process scores in conjunction with typical scores of performance (accuracy, total time, etc.) might enhance diagnostic prediction.
A third alternative in evaluating situations in which construct validity has not been fully supported is the possibility that the constructs are not independent from one another and thus could not be assessed as separate constructs. However, WM and IC have been consistently dissociated in children and adults on performance measures (Pennington and Ozonoff, 1998) and rating scales (Burgess et al., 1997; Gioia et al., 2001). In addition, although all IC measures were not more strongly related than WM measures, the lack of findings are likely attributed to differences in the assessment methods of rating scales compared to performance measures and the potential lack of specificity in isolating assessment of the two constructs rather than the lack of independence in the constructs (as discussed above). In addition, when the relationships between performance measures and rating scales are examined separately, some support is obtained for construct validity of WM and IC, more so among the behavior rating measures than the performance measures. Among the rating scales, the BRIEF Inhibit scale was more strongly correlated with the BRIEF BRI and the CPRS hyperactivity-impulsivity scale whereas the BRIEF Working Memory was more strongly correlated with the BRIEF MCI and CPRS inattention scale. Among the performance measures, the correlations between all performance measures of IC (homomethod-homotrait) were slightly higher with one another, than with the measures of WM (homomethod-heterotrait) suggesting that these constructs were also dissociated among the performance measures although to a less significant degree. Overall, findings from both behavior rating and performance measures suggest that the two constructs can be dissociated from one another.

Nonetheless, the independence of WM and IC does not discount that they are likely highly interconnected processes. Many recent models of EF have proposed that there is a
limited capacity, central pool of resources that underlies both WM and IC (e.g., Bjorklund & Harnishfeger, 1990; Cohen & Servan-Schreiber; 1992, Conway & Engle, 1994; Harnishfeger & Bjorklund, 1994; Kimberg & Farah, 1993) That is, both cognitive processes draw upon the same pool of resources; thus, demands for working memory interfere with inhibitory control and demands for inhibitory control disrupt the operation of working memory. Following from this perspective, one could view tasks that have traditionally been identified with one of these cognitive processes as actually demanding both cognitive processes, but perhaps different in the degree to which WM or IC is engaged. Therefore, the notion that there are ‘‘pure’’ tasks of either WM or IC may be misguided, as it is likely that such tasks differ in the relative contributions of each to the final behavioral outcome.

The interdependence of WM and IC are also likely greater within the developing brain. While many of the common EF measures have well documented utility in adult populations, it is yet unclear whether adult measures tap similar skills or utilize similar neuroanatomical regions in children. EF tests that assess fully developed EF skills in adults may engage different skills at different levels of development throughout childhood, especially in children with neurodevelopmental disorders whose core neuropsychological profile consists of delays or abnormal development of EF (Anderson, 1998). In addition, acquisition of EF skills undergoes rapid and variable change across childhood making it difficult to develop measures of EF that are appropriate for assessment across the development spectrum (Anderson, 2002). Several studies suggest there is a stage like developmental trajectory of EF skills during childhood (Becker, Isaac, & Hynd., 1987; Passler et al., 1985; Welsh, Pennington, & Groisser, 1991) that
coincides with maturational spurts in frontal lobe development (Bell & Fox, 1992; Levin et al., 1991; Thatcher, 1991, 1992; Welsh & Pennington, 1988). The studies by Levin et al. (1991), Welsh, Pennington, & Groisser (1991), and Anderson (1988) all suggest that EF skills undergo significant development between the ages of 3 and 12 years, especially in the years between 6 and 12, but they offer somewhat different models for which skills develop at different ages. Finally, Anderson (2002) provides data suggesting there is a temporary regression in EF performance around the ages of 11 and 12, particularly in areas of self-regulation and strategic decision-making. He suggests this regression may reflect a developmental transitional period between developmental phases of EF acquisition due to conflict from developing cognitive processes that lead to reintegration of EF skills. Thus, validating use of EF performance measures within a developmental framework is much more difficult than with traditional paradigms in adult populations, which is one of the central arguments that Anderson suggest that EF measures should be as specific in their assessment as possible in order to better disentangle developmental trajectories of different EF skills.

A different issue relevant to the assessment of EF performance of children is the potential moderating effect of IQ. It is generally acknowledged within the adult literature that measures of EF assess aspects of reasoning and cognition that are distinct from intelligence (Johnstone, Holland, & Larimore, 1997; Donders and Kirsch, 1991), and that IQ does not contribute substantially to understanding the impact of a lesion on higher cognitive ability, nor do lesions of the frontal lobes, or even full frontal lobectomies, necessarily impair IQ (Hebb, 1945; Brazzelli et al., 1994; Damasio and Anderson, 1993). In contrast, the relationship between IQ and EF in children is the subject of some debate
(Ardilla, et al., 2000; Arffa, et al., 1998; Mahone et al., 2002), and there is increasing evidence that IQ may play a moderating effect at higher levels of IQ (Baron, 2003). In a study of ADHD children, Mahone et al. (2002) found that IQ was significantly related to performance on EF measures, with IQ accounting for a greater portion of the variance than the ADHD diagnosis. Further, the authors reported that the impact of diagnosis was more salient for children with IQs within the average range, whereas differences between children with and without ADHD on performance measures of EF decreased for children with above average or superior IQ. These results are consistent with findings that the relationship between EF and IQ is stronger at the lower range of IQ (Duncan et al., 1995), and that children with superior IQ (above 135) outperform children with above average IQ (between 110 and 124) on EF variables (Arffra et al., 1998). Mahone et al. (2002) argued that IQ may serve as a marker for the integrity of prefrontal systems modulating executive control. These results are consistent with Denckla’s (1994) assertion that EF ability should be evaluated within the context of one’s overall intellectually ability as children with higher IQ may be able to perform adequately on performance measures of EF.

Despite the impact of IQ, EF tests are considered to measure constructs distinct from intelligence and there is ample data to support the dissociation of the two (Welsh and Pennington, 1998), although the dissociation is more supported in adults than children (Ardilla, Galeano, & Rosselli, 1998; Ardilla et al., 2000; Ardila, Pineda, & Rosselli, 2000; Boone, Ghaaffarian, & Lesser, 1993; Golden, 1981; Riccio, Hall, Morgan, & Hynd, 1994; Waldemann, Dickson, Monahan, & Kazelskis, 1992). Despite moderate correlations between IQ and EF, Baron (2003) states, “since high intelligence is not a
guarantee of flexible thinking, an extremely bright individual can demonstrate incapacitating cognitive rigidity and limiting rule-bound behavior contrary to the expectation of others. Similarly, lower intelligence does not dismiss the possibility of good common sense and creativity lead to effect ‘overachievement’ and an ability to conceptualize beyond the routine” (page 139). Distinguishing between crystallized and fluid intelligence is likely critical to disentangling the effects of IQ on EF performance (Duncan et al., 1995) as EF is more associated with measures of fluid than crystallized intelligence and performance aspects of intelligence testing that require novel planning and reasoning are more likely to correlate with EF (Denckla, 1996b). Barkley (1998) suggested that ADHD might affect IQ by impairing the implementation of skills (i.e. ineffective use of existing skills) and, to a smaller degree, the acquisition of intellectual skills.

Overall, this study supports the assertion by many investigators that assessment of EF requires assessment of both performance and behavior rating measures of EF (Baron, 2003; Burgess, 1997; Anderson, 2002). Results were consistent with the suggestions of many investigators that performance measures of EF are insufficient to discriminate ADHD children from controls and should not be used in isolation if ecological validity is desired. While the convergent validity of WM and IC across all measures was modest at best, there was a pattern of correlations within the behavior rating scales and performance measures to confirm that WM and IC are dissociable as independent constructs. In addition, the association of WM with symptoms of inattention and of IC with symptoms of hyperactivity-impulsivity provided strong evidence for divergent validity and further
support that EF is a valid and useful framework for understanding the behavioral characteristics of ADHD.

Strengths of this study include the well characterized control and diagnostic groups, the sampling of children across the middle school years, and the assessment of WM and IC constructs across a wide range of standardized rating scales and performance measures. The significant differences in IQ and age are a critical weakness of this study. Normal controls in this study had an average IQ within the above average range, which was significantly greater than the ADHD groups, both of which had average IQ’s within the average range that were not significantly different one another. In addition, IQ was significantly correlated with most of the measures, accounting for an average of 13% of the variance, with the highest correlations between IQ and the rating scale measures. IQ correlated more strongly with the D-KEFS performance measures than the unstandardized neurobehavioral performance measures. While IQ is a potential confound for this study given the strong correlations with the EF measures used, it is unclear whether IQ reduced the ability of this study to demonstrate differences between groups as one might have predicted greater group differences given the lower (within the average range) IQ of ADHD groups compared to the (above average) IQ for controls. It would seem reasonable to expect that the differences in IQ should have inflated between group differences in EF performance in this sample.

However, a larger confound for this study was the significant difference in age between the groups, which also demonstrated an interaction with IQ. Normal controls, with above average IQ, were younger than both ADHD groups. Similarly, children with ADHD-C/HI were younger and had a higher average IQ than children with ADHD-I,
although the differences were insignificant. In general IQ scores declined with increasing age. In addition, there was a significant effect of age on the BRIEF rating scales such that higher scores (more impaired) were associated with older children, which was likely attributed to the older age of the ADHD groups. The dual impact of higher IQ and lower scores on EF rating scale measures may have provided an advantage to younger children, allowing them to perform similarly to the older children with lower IQ and higher levels of EF dysfunction and ADHD symptomatology. Thus, the ability to detect differences between groups was significantly limited by the unequal distribution of diagnostic group and difference in IQ across the age groups.

Collection of more subjects to balance the age and IQ ranges across the groups would be extremely important in clarifying results from this study. In addition, inclusion of more ADHD-I and ADHD-HI children across the all age groups would assist in differentiating the effect of age from inattentive and hyperactive-impulsive symptoms on EF ability and enable exploration of the developmental trajectory of ADHD. Further, although gender was not a focus of this study and the groups were similar in the distribution of gender, it would be worthwhile to explore the impact of gender in the complex interaction of EF development and ADHD symptomatology. While the majority of developmental disorders, ADHD included, have a male bias, and most research indicates that boys and girls acquire EF skills at a similar rate during childhood (Becker et al., 1987; Chelune & Baer, 1986; Passler et al., 1985; Welsh et al., 1991), the small number of girls in this study limits the power of this study explore possible sex differences with the control and ADHD groups.
### APPENDIX A

**AGE CORRELATIONS**

Table 14. Correlations of Age with Rating Scales and Performance Measures

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>p</th>
<th>r²</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ</td>
<td>-0.34</td>
<td>0.01</td>
<td>0.12</td>
<td>0.73</td>
</tr>
<tr>
<td>BRIEF Inhibit</td>
<td>0.28</td>
<td>0.02</td>
<td>0.08</td>
<td>0.58</td>
</tr>
<tr>
<td>BRIEF BRI</td>
<td>0.26</td>
<td>0.03</td>
<td>0.07</td>
<td>0.53</td>
</tr>
<tr>
<td>BRIEF WM</td>
<td>0.27</td>
<td>0.02</td>
<td>0.07</td>
<td>0.55</td>
</tr>
<tr>
<td>BRIEF MCI</td>
<td>0.27</td>
<td>0.02</td>
<td>0.07</td>
<td>0.56</td>
</tr>
<tr>
<td>CPRS - I</td>
<td>0.23</td>
<td>0.05</td>
<td>0.05</td>
<td>0.47</td>
</tr>
<tr>
<td>CPRS - H</td>
<td>0.10</td>
<td>0.41</td>
<td>0.01</td>
<td>0.20</td>
</tr>
<tr>
<td>Trails 4</td>
<td>-0.23</td>
<td>0.06</td>
<td>0.05</td>
<td>0.46</td>
</tr>
<tr>
<td>CF Switching</td>
<td>0.18</td>
<td>0.13</td>
<td>0.03</td>
<td>0.37</td>
</tr>
<tr>
<td>CWI Inhibition</td>
<td>-0.06</td>
<td>0.64</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>Tower</td>
<td>-0.13</td>
<td>0.28</td>
<td>0.02</td>
<td>0.26</td>
</tr>
<tr>
<td>Span Backwards</td>
<td>-0.12</td>
<td>0.31</td>
<td>0.02</td>
<td>0.25</td>
</tr>
<tr>
<td>Conflicting Motor Resp.</td>
<td>0.39</td>
<td>0.00</td>
<td>0.15</td>
<td>0.84</td>
</tr>
<tr>
<td>Lateral Gaze</td>
<td>0.31</td>
<td>0.01</td>
<td>0.10</td>
<td>0.66</td>
</tr>
<tr>
<td>SOP Between Errors</td>
<td>-0.34</td>
<td>0.01</td>
<td>0.12</td>
<td>0.73</td>
</tr>
<tr>
<td>Simple GNG</td>
<td>-0.47</td>
<td>0.00</td>
<td>0.22</td>
<td>1.06</td>
</tr>
<tr>
<td>Counting GNG</td>
<td>0.21</td>
<td>0.11</td>
<td>0.05</td>
<td>0.43</td>
</tr>
</tbody>
</table>

**NOTE:** * Indicates significance at \( p < 0.01 \) and \( d > 0.5 \)

FSIQ = WISC Full Scale IQ; BRIEF = Behavior Rating Inventory of Executive Function; Inhibit = Inhibition Scale; BRI = Behavioral Regulation Index; WM = Working Memory Scale; MCI = Metacognition Index; CPRS – I = Conners Parent Rating Scale DSM Inattentive (Scale L); CPRS - HI = Conners Parent Rating Scale DSM Hyperactive-Impulsive (Scale M); Trails 4 = D-KEFS Trails Condition 4 Total Time Scaled Score; CF Switching = D-KEFS Category Switching Total Scaled Score; CWI Inhibition = D-KEFS Color Word Interference Inhibition Total Time Scaled Score; Tower Achievement = D-KEFS Tower Total Achievement Scaled Score; Span Backwards = WISC Spatial Span Backwards Scaled Score; Conflicting Motor Resp. = Conflicting Motor Response z-score; Lateral Gaze = Lateral Gaze Total Time z-score; SOP Between Errors = Self-Ordered Pointing Between Errors z-score; Simple Go No-Go = Simple Go No-Go Commissions z-score; Counting Go No-Go = Counting Go No-Go Omissions z-score.
# APPENDIX B

## IQ CORRELATIONS

Table 15. Correlations of IQ with Rating Scales and Performance Measures

<table>
<thead>
<tr>
<th></th>
<th>$r$</th>
<th>$p$</th>
<th>$r^2$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIEF Inhibit</td>
<td>-0.33</td>
<td>0.00</td>
<td>0.11</td>
<td>0.70</td>
</tr>
<tr>
<td>BRIEF BRI</td>
<td>-0.37</td>
<td>0.00</td>
<td>0.13</td>
<td>0.79</td>
</tr>
<tr>
<td>BRIEF WM</td>
<td>-0.46</td>
<td>0.00</td>
<td>0.21</td>
<td>1.02</td>
</tr>
<tr>
<td>BRIEF MCI</td>
<td>-0.42</td>
<td>0.00</td>
<td>0.18</td>
<td>0.92</td>
</tr>
<tr>
<td>CPRS - I</td>
<td>-0.40</td>
<td>0.00</td>
<td>0.16</td>
<td>0.87</td>
</tr>
<tr>
<td>CPRS - H</td>
<td>-0.28</td>
<td>0.02</td>
<td>0.08</td>
<td>0.57</td>
</tr>
<tr>
<td>Trails 4</td>
<td>-0.52</td>
<td>0.00</td>
<td>0.27</td>
<td>1.22</td>
</tr>
<tr>
<td>CF Switching</td>
<td>0.46</td>
<td>0.00</td>
<td>0.21</td>
<td>1.04</td>
</tr>
<tr>
<td>CWI Inhibition</td>
<td>-0.35</td>
<td>0.00</td>
<td>0.12</td>
<td>0.74</td>
</tr>
<tr>
<td>Tower</td>
<td>0.23</td>
<td>0.06</td>
<td>0.05</td>
<td>0.47</td>
</tr>
<tr>
<td>Span Backwards</td>
<td>0.39</td>
<td>0.00</td>
<td>0.16</td>
<td>0.86</td>
</tr>
<tr>
<td>Conflicting Motor Resp.</td>
<td>0.41</td>
<td>0.00</td>
<td>0.17</td>
<td>0.90</td>
</tr>
<tr>
<td>Lateral Gaze</td>
<td>0.23</td>
<td>0.06</td>
<td>0.05</td>
<td>0.48</td>
</tr>
<tr>
<td>SOP Between Errors</td>
<td>-0.32</td>
<td>0.02</td>
<td>0.10</td>
<td>0.67</td>
</tr>
<tr>
<td>Simple GNG</td>
<td>-0.12</td>
<td>0.37</td>
<td>0.01</td>
<td>0.23</td>
</tr>
<tr>
<td>Counting GNG</td>
<td>0.24</td>
<td>0.07</td>
<td>0.06</td>
<td>0.50</td>
</tr>
</tbody>
</table>

| Average effect size      | 0.75 |
| Average effect size for Rating Scale correlations | 0.81 |
| Average effect size for D-KEFS correlations | 0.87 |
| Average effect size for Neurobehavioral EF correlations | 0.55 |

**NOTE:** * Indicates significance at $p < 0.01$ and $d > 0.5$; BRIEF = Behavior Rating Inventory of Executive Function; Inhibit = Inhibition Scale; BRI = Behavioral Regulation Index; WM = Working Memory Scale; MCI = Metacognition Index; CPRS - I = Conners Parent Rating Scale DSM Inattentive (Scale L); CPRS HI = Conners Parent Rating Scale DSM Hyperactive-Impulsive (Scale M); Trails 4 = D-KEFS Trails Condition 4 Total Time Scaled Score; CF Switching = D-KEFS Category Switching Total Scaled Score; CWI Inhibition = D-KEFS Color Word Interference Inhibition Total Time Scaled Score; Tower Achievement = D-KEFS Tower Total Achievement Scaled Score; Span Backwards = WISC Spatial Span Backwards Scaled Score; Conflicting Motor = Conflicting Motor Response z-score; Lateral Gaze = Lateral Gaze Total Time z-score; SOP Between Errors = Self-Ordered Pointing Between Errors z-score; Simple Go No-Go = Simple Go No-Go Commissions z-score; Counting Go No-Go = Counting Go No-Go Omissions z-score.
### APPENDIX C
CORRELATIONS AMONG ALL MEASURES

Table 16. Correlations of BRIEF and Performance Measures

<table>
<thead>
<tr>
<th></th>
<th>BRIEF Inhibit</th>
<th>BRIEF BRI</th>
<th>BRIEF WM</th>
<th>BRIEF MCI</th>
<th>Trails 4</th>
<th>CF Switching</th>
<th>CWI Inhibition</th>
<th>Tower</th>
<th>Span Backwards</th>
<th>Conflicting Motor Resp.</th>
<th>Lateral Gaze</th>
<th>SOP Between Errors</th>
<th>Simple Go-No-Go</th>
<th>Counting Go-No-Go</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIEF Inhibit</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRIEF BRI</td>
<td>***0.94</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRIEF WM</td>
<td>***0.64</td>
<td>***0.65</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRIEF MCI</td>
<td>***0.73</td>
<td>***0.72</td>
<td>***0.95</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails 4</td>
<td>-0.02</td>
<td>0.12</td>
<td>-0.14</td>
<td>-0.19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF Switching</td>
<td>-0.08</td>
<td>-0.06</td>
<td>0.02</td>
<td>0.04</td>
<td>-0.21</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CWI Inhibition</td>
<td>0.06</td>
<td>0.27</td>
<td>0.11</td>
<td>0.1</td>
<td>0.24</td>
<td>-0.22</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tower</td>
<td>-0.14</td>
<td>-0.19</td>
<td>*-0.32</td>
<td>-0.29</td>
<td>0.09</td>
<td>0.04</td>
<td>-0.09</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Span Backwards</td>
<td>-0.07</td>
<td>-0.08</td>
<td>0.03</td>
<td>0.08</td>
<td>-0.02</td>
<td>-0.25</td>
<td>-0.24</td>
<td>-0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflicting Motor Resp.</td>
<td>-0.29</td>
<td>-0.29</td>
<td>-0.17</td>
<td>-0.29</td>
<td>0.17</td>
<td>-0.12</td>
<td>-0.29</td>
<td>0.1</td>
<td>0.19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral Gaze</td>
<td>***-0.46</td>
<td>***-0.6</td>
<td>***-0.59</td>
<td>***-0.62</td>
<td>-0.01</td>
<td>0.07</td>
<td>-0.3</td>
<td>0.09</td>
<td>-0.12</td>
<td>*0.31</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOP Between Errors</td>
<td>0.14</td>
<td>0.21</td>
<td>*0.35</td>
<td>**0.39</td>
<td>-0.13</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.2</td>
<td>0.04</td>
<td>-0.23</td>
<td>**-0.4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple Go-No-Go</td>
<td>0.25</td>
<td>0.23</td>
<td>*0.33</td>
<td>*0.37</td>
<td>-0.28</td>
<td>-0.03</td>
<td>0.21</td>
<td>-0.27</td>
<td>0.27</td>
<td>-0.3</td>
<td>-0.25</td>
<td>0.21</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Counting Go-No-Go</td>
<td>-0.29</td>
<td>-0.27</td>
<td>***-0.45</td>
<td>***-0.48</td>
<td>0.02</td>
<td>0.1</td>
<td>-0.03</td>
<td>*0.38</td>
<td>-0.22</td>
<td>0.19</td>
<td>0.24</td>
<td>-0.27</td>
<td>***-0.59</td>
<td>1</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES


BIOGRAPHICAL SKETCH

Christopher William Loftis, M.S. was born and raised in Southern California. He received a B.A. with high honors in psychology, in 1995, from the University of California, San Diego (UCSD), where he was involved with research of behavioral inhibition, linguistics, and child traumatic brain injury. After graduating from UCSD, he worked as a study coordinator and lab manager on several grants investigating cognitive impairments associated with aging, Alzheimer’s Disease, Parkinson’s Disease, Lewy Body Variant, and alcohol abuse.

He received a master’s degree in clinical psychology in 1999, from the Department of Clinical and Health Psychology at the University of Florida, for his functional Magnetic Resonance Imaging (fMRI) investigation of the role of the hippocampus in encoding and retrieval memory processes, supported by funding from the NIH National Research Service Award. From July 2002 to June 2003, he completed a clinical internship in pediatric neuropsychology and behavior psychology at the Kennedy Krieger Institute (KKI) in Baltimore. He then joined the Department of Developmental Cognitive Neurology at KKI when he obtained a Dissertation Grant from the Maternal and Child Health Bureau to complete his doctoral dissertation on the ecological validity of executive function measures in Attention-Deficit Hyperactivity Disorder. He anticipates graduation in August 2004.

In addition to his academic pursuits, Mr. Loftis is actively involved in national professional organizations and has served as chair of the Advocacy Committee for
American Psychological Association of Graduate Students (APAGS), as Chair of the APAGS, and as Chair of the Association for Neuropsychology Students in Training, a subcommittee of the APA Division 40 (Clinical Neuropsychology). He is currently serving on the APA Board of Directors and Council of Representatives.

He has received several awards including the APA Board of Educational Affairs Special Award, Health Professions Dean Scholar Award, Grace Winslow Shands Hospital Auxiliary Scholarship, Army Aviation Association of America Scholarship Foundation, and the Charlotte Liberty Scholarship.