DOES DIRECT ASSESSMENT OF EXECUTIVE FUNCTIONS PREDICT OBSERVED EXECUTIVE DYSFUNCTION IN A CLINICAL SAMPLE OF CHILDREN?: A COMPARISON OF LATENT ASSESSMENT VARIABLES WITH OBSERVER RATINGS

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

DEPARTMENT OF PSYCHOLOGY AND PHILOSOPHY

COLLEGE OF ARTS AND SCIENCES

BY

JOSEPH JEANE-LEEMAN, M.S.

DENTON, TX

DECEMBER 2015

ACKNOWLEDGEMENTS

To my lovely wife, Stephanie, who redefines patience. You sacrificed so much to help me make it through the last nine years: I'm all yours now! To my beautiful daughter who has been alive only slightly longer than I have been in school. Lily, I can finally answer the question you asked so many times: "When are you going to finish that dumb paper?" To my Dad, who helped me understand that *The Noble Path* was a good alternative to the path of despondence. And to my Aunt Marion: what can I say dearest Aunt? My title should forever be Dr. Jeane-Leeman, *sponsored by Marion Leeman*, because without you, the <u>goal would have been unreachable</u>. To my extended family, Bryan and Dana; thank you for letting me know there will be life after dissertation. I love you all.

I would be remiss if I did not those who were directly involved in the effort to finish this journey. I owe a great deal of my success to the excellent internship supervision of Dr. Ensign, and the editing of Dr. Lowther. To those professors who graciously accepted a role on my committee, I offer special thanks: Dr. DeOrnellas, for her ability to listen patiently; Dr. Maricle, for helping me overcome my fears; Dr. Miller, for advancing the practice of school psychology as well as my understanding of neuropsychology; and to Dr. Asbury, for introducing me to Andy Field. Hey, Doc: I finally stopped running numbers!!

ABSTRACT

JOSEPH JEANE-LEEMAN M.S.

DOES DIRECT ASSESSMENT OF EXECUTIVE FUNCTIONS PREDICT OBSERVED EXECUTIVE DYSFUNCTION IN A CLINICAL SAMPLE OF CHILDREN?: A COMPARISON OF LATENT ASSESSMENT VARIABLES WITH OBSERVER RATINGS

DECEMBER 2015

Executive functions are processes which allow an individual to regulate and direct emotional, cognitive, perceptive, and motor processes to engage in purposeful, goal directed behavior (McCloskey, Perkins, & Van Divner, 2006). Deficits in executive processes are also central to many clinical childhood disorders and difficulties with academic and social learning (Latzman, Elkovitch, Young, & Clark, 2010; Mattison & Mayes, 2012). Despite professional agreement on the importance of executive functioning, there has been considerable debate as to how this construct should be conceptualized and evaluated (Meltzer & Krishman, 2007). The purpose of this research study was to explore the relationship between measurements of executive processes and observer ratings of executive dysfunction in everyday activities. Data from formal assessment instruments from a clinical sample of children 8 to 16 years of age were subjected to an exploratory factor analysis. Latent variables derived from this analysis were used to predict observer ratings of adequate versus impaired executive processes in the sample. Subtests used in analysis of formal assessment were drawn from test batteries commonly used in assessment of children: the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001a), A Developmental Neuropsychological Battery, Second Edition (NEPSY-II; Korkman, Kirk, & Kemp, 2007a), and the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition Normative Update (WJIII COG NU; Woodcock, McGrew, & Mather, 2007). The parent form of the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia, Isquith, Guy & Kenworthy, 2000a) was used to determine the level of observed executive dysfunction in everyday activities. Results suggest that factors in the clinical sample approximated, but were less differentiated than, current models of executive function suggest. Factors consisted of tasks, which facilitated performance on executive tasks, as well as, tasks that directly measured executive processes. Latent variables derived from the factors were not adequate predictors of observer ratings of executive dysfunction in the overall sample; however, a factor comprising concept formation and working memory was a weak predictor of group membership in Caucasians. The implications of the research for practitioners assessing executive processes in referred children are discussed.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Executive Function (EF) comprises the information processing skills chiefly associated with the pre-frontal cortex (PFC) of the brain that mediate cognitive and emotional resources in goal directed behavior (Royall et al., 2002). These metacognitive skills serve to activate and inhibit behavioral sequences, while sustaining orientation towards a specific goal (Miyake & Friedman, 2012). The importance of EF to learning across academic domains has been validated from primary to secondary levels (Bull & Scerif, 2001; Latzman et al., 2010). Differences in executive skills are primarily a function of biological and environmental factors that affect neurological processes of cognitive and emotional regulation. Deficits in various aspects of cognitive control are most pronounced in cases of traumatic brain injury, but differing degrees of impairment in executive skills such as inhibition, mental flexibility, and planning, are associated with various neurologically-based disabilities, such as high functioning autism (McCrimmon, Schwean, Saklofske, Montgomery, & Brady, 2012), and attention deficit hyperactivity disorder (Dickstein, Bannon, Castellanos, & Milham, 2006).

Despite agreement that EF is central to efficient cognitive processing, there has been considerable professional debate as to how EFs should be conceptualized and evaluated (Meltzer & Krishman, 2007). Although many neuropsychological assessment tasks are adapted from instruments with proven empirical validity in detecting executive dysfunction, they are not bound by a cohesive theoretical model (Homack, Lee & Riccio, 2005). Attempts to discern the factor structure of executive processes have been elusive due to the number of instruments used in research as well as the breadth of proposed constructs (Packwood, Hodgetts, & Tremblay, 2011).

The assessment of executive processes in children has proven even more problematic due to differing developmental trajectories resulting in differences in factor structure of assessments found in different age groups (Brookshire, Levin, & Song, 2004; Denham, Warren-Khot, Bassett, Wyatt, & Perna, 2012). Some clinicians have further argued that structured assessment does not capture the everyday manifestation of executive processes and is therefore less ecologically valid than rating scales (Gioia, Isquith, Retzlaff, & Espy, 2002). The established importance of executive processes in social and academic learning has driven the need for further exploration of executive dysfunction in clinical samples to provide better assessment and intervention. This study will add to the increasing body of research on the structure of EF in a clinical sample of children using assessment tasks commonly employed by clinicians in schools. This research will also aide practitioners in better understanding how combinations of common assessment instruments reflect observed executive deficits in children.

Brief Literature Review

History

The concept of metacognitive processes that govern thought and behavior has its roots in the philosophical concept of the mind, which has occupied philosophers from

Aristotle to Descartes (Wenzel, 2006). As philosophy gave way to science and a dualistic view of mind and body gave way to a biological view of thought and behavior, the neurosciences were born. In the late 1800's the rise of experimental psychology led to empirical data on brain organization and function. It was these explorations that resulted in William Wundt positing a central role to the frontal cortex and providing the first definition of executive processes as the conscious processing of sensory data through willed, focused attention (Wundt, 1910).

The understanding of how EF associated with prefrontal areas of the brain interacted with lower cognitive and neurological processes increased in the early 20^{th} century. Through his study of brain injured and aphasic patients, Soviet physician Alexander Luria developed an empirically based structural model of brain function that endures today (Languis & Miller, 1992). Luria further defined facets of executive processes by describing the role of the PFC in initiation and inhibition of a conditioned response (Banich, 2009). Modern brain imaging techniques have largely supported Luria's model revealing that the PFC has diffuse connections to primary and secondary brain regions that form independent regulatory circuits much as Luria predicted (Royall et al., 2002). As imaging techniques have become more advanced, executive processes have been increasingly conceptualized as manifesting through an interplay of cortical networks and variances in cortical activation depending on diverse factors such as age, pathology, sensory modality, and task demands (Carpenter, Just, & Reichle, 2000).

With a greater understanding of neurological dysfunction came a greater emphasis on assessing and quantifying observed deficits. British neurologist Ward Halstead and his colleague, Hal Reitan developed an assessment battery that would become the primary paradigm for neuropsychological assessment (Reitan & Wolfson, 1993). Other assessments such as trail making tasks, the Stroop color word task, category tests, and verbal fluency tasks (Mitrushina, Boone, Razani, & D'Elia, 2005) became commonly used to detect neurological processing difficulties. In recent years, the validity of some of these tasks, such as the Wisconsin Card Sorting Task, has been validated by brain imaging techniques that reveal diffuse frontal activation in problem solving. With further emphasis on the processes of task completion, the usage of these tests shifted from diagnosing pathology to analyzing strengths and weaknesses in neuropsychological functioning that can explain differences in the acquisition of skills (Miller, 2013).

Theoretical Models

The theoretical view of EF largely depends on whether it is viewed as a unitary construct, which cannot be subdivided into discrete processes, or if it is viewed as a set of related but distinctly separate skills. Theories which view EF as a single unitary construct have historically thought of the PFC as a supervisory executive system governing attention (Baddeley, 2000). Others equate it with overall intelligence or a central aspect of cognitive theory such as fluid reasoning (Blair, 2006). Most theories postulate multiple distinct, but related, processes associated with goal direction and behavioral regulation. In recent years, some researchers have attempted to correlate proposed constructs with neuroimaging, identifying specific pathways between prefrontal and limbic system components that consistently regulate skills such as response inhibition (Aron, 2008). However, other studies testing a broader range of proposed skills have found activation in similar prefrontal regions regardless of the skill assessed (Duncan & Owen, 2000). In some models, factors such as response inhibition, planning, and cognitive flexibility are largely based on self-regulation skills that have shown observable deficits in clinical populations (Purdy, 2005). Other models base their factors on differences in cognitive and behavioral regulation or what has been termed "hot" and "cool" EF (Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009). The school neuropsychological processing model (SNP) views EF as a broad band classification encompassing narrow band skills such as inhibition, fluency and problem solving (Miller, 2013). Whereas the SNP places working memory as a facilitator of cognitive processes, other research has shown that tests of working memory and executive processes measure the same underlying construct (Collette & Van der Linden, 2002; Donders, Denbraber, & Vos, 2010). Some models propose a complex, task dependent view of executive control (McCloskey, et al., 2006) while other models attempt a parsimonious solution to capturing executive processes (Miyake et al., 2000). Duplication of models in childhood populations are inconsistent as neurological changes invoke different sets of regulatory processes at different age levels (Meltzer & Krishman, 2007). In general, the younger the age group assessed, the more difficult it becomes to separate factors (Lee, Bull, & Ho, 2013).

Neuroanatomy of Executive Processes

The PFC represents Luria's tertiary area of the brain, that which controls initiation and inhibition of behavioral sequences necessary for goal attainment (McCloskey et al., 2006). Distinct areas of the PFC have different roles in regulating behavior. Major functional circuits include the dorsolateral prefrontal cortex (dlPFC), involved in regulating cognitive and conceptual processes; the orbitofrontal cortex (OFC), involved in emotional regulation; and the ventral medial prefrontal cortex (vmPFC), which is key in emotional decision-making and motivation to action (Cummings & Miller, 2007). Additionally, the inferior frontal cortex (IFC) has been implicated in the inhibition of stopping an action that is in process (Aron, 2008). The anterior cingulate (ACC) is a critical component to the suspension of habitual response patterns and in regulating the level of activation of the OFC and dlPFC when existing strategies are unsuccessful (De Pisapia, Slomski, & Braver, 2007). All of these cortical structures connect extensively in a feedback loop with the limbic system and other lower brain units much as Luria envisioned. The logical conclusion is that poor performance on executive tasks may not be due to executive dysfunction but instead indicate inefficiency in more basic cognitive processes. This is why it is important to test core executive skills in different modalities and under different conditions.

Development

Difficulties with assessment of executive processes in children are a reflection of cortical development and the emergence of new skills. From ages 3 to 6, children experience marked spurts of PFC dendritic growth and cortical volume, with myelination of these connections in secondary areas peaking by age 7, and in the PFC during adolescence (Rathbone et al., 2011). Skills do not always develop on a linear timetable. Selective attention skills show rapid development from age 3 to 6 with another rapid improvement between 8 and 10 years of age. Self-regulation skills are evidenced in infants as young as 14 months and again show rapid development as evidenced on inhibitory tasks between 8 and 10 years of age (Spencer-Smith & Anderson, 2009). Maturation of EF, in concert with increased processing speed, explains significant variance in task performance between age groups (Illig, 1998).

Genetic and Environmental Impact

Twin studies have revealed a significant influence of hereditability on the efficiency of executive processes (Friedman, et al., 2008) . In fact, as a unitary construct, executive functioning shows a higher genetic influence than overall cognitive ability, which exhibits considerably more environmental influence. Recently research has revealed high levels of cortisol present in children due to adverse environmental affects specifically due to socioeconomic factors and lack of maternal nurturance (Meaney, 2010). Cortisol increases emotional reactivity and has been implicated in disorders that affect behavioral regulation. These effects are exacerbated by poor parental care or trauma which significantly impacts early development of executive processes key to learning (Blair et al., 2011). Studies on children who are at-risk due to socioeconomic variables such as prenatal risk factors, premature birth, lack of nutrition, high parental stress, and lack of stimulation, exhibit decreased density in multiple brain areas including the OFC and anterior cingulate (Hackman, Farah, & Meaney, 2010).

Executive Dysfunction and Disorders of Childhood

Executive functioning has been found to play a significant role in academic achievement with specific factors varying in importance across tasks and age levels (Latzman et al., 2010). Deficits in executive attention control in at-risk students may explain negative academic experiences that show declining levels of motivation over time (Howse, Lange, Farran, & Boyles, 2003). These deficits in executive attention and control may also explain the higher prevalence of ADHD among low socioeconomic groups (Ardila, Pineda, & Rosselli, 2000). Self-regulation in children diagnosed with ADHD has been conclusively tied to hypo-activation of the dlPFC and anterior cingulate regions of the frontal cortex (Dickstein et al., 2006).

Executive dysfunction is also linked to Autism where deficits in working memory, planning, and switching rule sets occur in less structured tasks that require formation of strategies for successful task completion (Hill, 2004). Traumatic brain injury (TBI) almost always involves executive dysfunction and severe TBI can impair regulation and goal setting abilities a decade after injury (Beauchamp et al., 2011). Deficits in executive processes such as working memory have also been implicated in the development of specific learning disabilities in math and reading (García-Madruga, Vila, Gómez-Veiga, Duque, & Elosúa, 2014; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011).

Assessment of Executive Functions

Historically, specific formal assessment tasks have exhibited empirical validity in the detection of significant executive dysfunction (Mitrushina et al., 2005). Although these instruments have proven sensitive to significant impairments such as found in TBI patients, results in identifying differences in mild to moderate impairments found in childhood disorders have been inconsistent (Mahone et al., 2002). In addition to the aforementioned disagreement on definitions of executive processes, other confounding factors include a lack of age and gender sensitivity, the necessity of task novelty, which limits test-retest reliability, and the structured assessment setting, which fails to capture how skills are used in the environment. There is also the problem of task impurity, meaning it is difficult to surmise if one is measuring deficits in the executive processes or in the skills these processes regulate (Jurado & Rosselli, 2007). To better address "real-world" functioning, rating scales such as the Behavior Rating Inventory of Executive Functioning (BRIEF), have been developed to measure observer ratings of behaviors associated with executive dysfunction (Gioia, Isquith, Guy, & Kenworthy, 2000b).

In recent years, the accuracy of formal assessment of executive functioning for diagnosis and intervention has been questioned. Some research has found that in comparison to rating scales of EF, direct assessment, at best, measures a different construct and at worse, does not accurately reflect deficits observed in the natural environment (Toplak, West, & Stanovich, 2013). Others argue that assessments of EF are closely tied to cognitive measures and are almost unitary with measures of cognitive processes (Floyd, Bergeron, Hamilton, & Parra, 2010). Some evidence has been provided that executive rating scales correlate well with established scales designed to measure emotional and behavioral disorders and thus are more an indication of pathology than neurological dysfunction (Bishop, 2010). This lack of agreement is further evidence that the conceptualization of executive processes is disturbingly disparate in psychological practice.

The central questions that arises from a review of current literature revolves around determining the constructs of executive processes that arise from formal assessments in children and how these constructs predict significant levels of observed EF in children with disabilities. This research seeks to find that common ground by examining both the factor structure of assessed EFs in a clinical sample of school aged children, and how those factors reflect clear differences between children who exhibit significant levels of observed executive dysfunction and those that do not. To accomplish this goal, this study examined the most commonly used and empirically validated instruments for assessment of EFs in children, as well as, a widely used observer rating of EF. The primary formal assessment instruments examined in this study were the Delis-Kaplan Executive Function System (D-KEFS; Delis et al., 2001a), the Developmental Neuropsychological Assessment Battery, Second Edition (NEPSY-II; Korkman, et. al, 2007a), and the Woodcock Johnson Tests of Cognitive Ability, Third Edition, Normative Update (WJIII COG NU; Woodcock, et al., 2007). The primary measure for observed EF was the parent version of the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia et al., 2000a).

Purpose and Relevance of the Research

The need for clinicians to find common ground in the definition and assessment of EFs should be a primary focus of research. Neurological impairments are the heart of many childhood disorders of cognitive and emotional regulation and these children are often misdiagnosed and mislabeled, leading to inappropriate and ineffective intervention (Miller, 2013). Children with executive dysfunction are often at risk for being labeled as "lazy" and the lack of understanding of how the brain develops after long term stress or trauma is sadly misunderstood in education (Meltzer & Krishman, 2007). The first step to a greater understanding is to begin to see a convergence of agreed upon facets of executive control central to learning, to identify instruments most effective in assessing these areas, and finally to explore at what level test results reflect observed deficits in self-regulation. This research sought to explore the factor structure of some of the most commonly used instruments in child populations that reflect the efficacy of executive processes. These latent variables representing formal assessment data were then analyzed to see what extent they were able to correctly predict significant observed executive dysfunction in the clinical sample.

Summary

The review of relevant research and the statistical analyses proposed led to the following questions:

1. Is there a discernible factor structure of EF in a clinical sample of children that emerges from commonly used assessment instruments having empirical validity in detecting executive dysfunction. If so, how many factors best describe the data and do these factors represent similar constructs reported in research on adult and child populations?
2. Can the latent variables derived from factors in the first analysis predict observer reported dysfunction in cognitive executive processes such as working memory, organization, planning and self-monitoring?

In the first research question, this author hypothesized that the results of the EFA would reveal more than one factor of EFs in the clinical sample. Further, the constructs that emerged would most likely represent some combination of inhibition, cognitive flexibility, selective attention, and working memory (Miyake et al., 2000). Based on previous research with children, it is possible that less factors or, highly correlated factors, may emerge.

In the second research question, this author hypothesized that the factors that emerged would accurately reflect observations of cognitive processing difficulties due to executive dysfunction to an extent greater than chance. The latent variables derived from the factor analysis would reflect and classify observer ratings of significant and non-significant levels of executive dysfunction as established by observational data obtained from the parent form of the BRIEF rating scales.

This research addresses the continued debate on the definition and assessment of specific EFs due in large part to a number of factors: the conceptualization of executive control and assessment, confounding factors in development, the proliferation of more established cognitive assessment theories and the lack of correlation between observed and assessed deficits in self-regulation. The assessment instruments analyzed in this study are commonly used in assessment of children and adolescents in the school and clinical setting. It is important to ascertain if the current methods of determining problems with executive processes are sensitive to such deficits and if they yield consistent results.

CHAPTER II

REVIEW OF LITERATURE

The following section is an in-depth review of the literature and research on EF. A historical perspective is presented highlighting important developments that have shaped current theoretical models and assessment practices. Evidence from neuroanatomical, neurochemical, and genetic research is presented to support the constructs being studied. Research is presented on the importance of these executive constructs to success in academic learning and in common disorders diagnosed in children. Current issues with theoretical approaches to conceptualizing and measuring executive processes are also discussed with an emphasis on the development, use and efficacy of commonly used, empirically validated assessment measures. Differences based on theory, and developmental factors are discussed. The section concludes with a discussion of how the current research might assist in better understanding the predictive validity of assessment in child clinical populations.

Defining Executive Functions

Executive function (EF) comprises the information processing skills chiefly associated with the PFC that mediate cognitive and emotional resources in goal directed behavior (Royall et al., 2002). When such functions are impaired or not effectively used, an individual's ability to regulate the complex interactions between perception, emotion, thought, and action is compromised resulting in serious difficulties in academic and social functioning (Aron, 2008; McCloskey, Hewitt, Henzel, & Eusebio, 2009). EFs have been defined to encompass such metacognitive processes as organizing, planning, initiating, and inhibiting behaviors while monitoring progress towards attaining immediate and future goals (McCloskey et al., 2006). Although few clinicians argue the importance of these self-regulatory processes, the concept of defining and integrating them into a cohesive model is so complex that current research reflects widespread disagreement on the conceptualization, primary constructs, and appropriate assessment of executive processes. For instance, in a meta-analysis of research articles on assessment of executive constructs, Packwood et al. (2011) found over 98 different assessment tasks used to theorize at least 18 different constructs of executive processes. The roots of these disagreements can be found in the historical beginnings and development of the field of neuropsychology.

History of Neuropsychology

In the 1600's, Renee Descartes expressed the dominant view of the mind as an indivisible construct of pure reason separate from the physical body (Brown, 2001). Descartes was describing the dualist philosophical view of mind and body, a concept which had endured dating back to Aristotle (Wenzel, 2006). This viewpoint encouraged the treatment of the mind as a homunculus that exerted executive control over lower processes. Benedict Spinoza, a contemporary of Descartes, was one of those who challenged this dominant dualist view of mind and body. Spinoza believed one's environment triggered emotional drives which initiated behaviors aimed at self-preservation while conscious thought reflected one's attempt to rationalize and understand these actions (Schmitter, 2014). The monist view emphasized observation and objective data collection over philosophical discussion and de-mystified the concept of the mind, allowing for the beginnings of empirical neurological research.

The 1700's to 1800's: The Localization Debate

By 1879, when William Wundt opened the first laboratory for the experimental study of psychology in Leipzig, Germany, the dominant theoretical paradigms for the study of brain function were established (Plucker, 2014). Drawings of the overall structure of the hemispheres accurately depicted brain structure, and there was little disagreement that the cerebral cortex was the seat of intelligence (Benton, 2000). Experimentation was centered on two theoretical paradigms. The first, localization theory, postulated that specific functions are inherent in specific geographic areas of the brain, and the second, that skills rose from the interaction of different brain regions. The first can be credited to physician Franz Joseph Gall, who is best remembered for his failed theory of phrenology stating that intellect and personality could be derived from the size and shape of the skull. The shape of the skull was influenced by the size and shape of the brain, which Gall believed to be made up of many separate organs each possessing localized control over a specific aspect of perception, movement, and thought. Gall's suppositions, published in the early 1820's, were at odds with the dominant paradigm pioneered by Jeane Marie Pierre Flourens, who reasoned that higher functions such as memory and attention arose from the unified interplay of cortical activity. Gall's theory of functional localization gained more prominence due to the findings of physicians such as Paul Broca and Karl Wernicke, who identified specific brain areas involved in speech production (Broca, 2011; Müller & Knight, 2006), and David Ferrier, who mapped areas governing specific motor movements through electro-stimulation of animal brains (Sabbatini, 2003). Research in neuropsychology began to center upon defining the specific function of each geographic area of the brain.

Early Conceptualization of Frontal Lobe Functioning

Proponents of localization theory had little success in explaining more diffuse functions such as memory and emotional regulation and, as a result, the study of the frontal lobes, especially those of the anterior frontal regions, received sparse attention (Benton, 2000). What was known from case reports was that damage in the anterior frontal lobes resulted in diverse changes in personality that were inconsistent and could not easily be translated into scientific terms. For instance, in the mid 1800's, a condition known as *witzelsucht*, or the tendency to joke excessively, was described in patients with frontal injuries. Study of the specific function of the PFC was sporadic until a famous case of personality change after injury was published. The injured man was Phineas Gage, a railroad worker in the mid-19th century, whose bilateral vmPFC was obliterated when a steel rod was propelled by an explosion through the front of his skull (Damasio, 1994). Surprisingly, Gage not only lived but experienced no noticeable changes in observed cognitive and motor functions such as memory, orientation, or motor competency necessary for employability. However, in a paper published by his physician, James Harlow, thirteen years after the injury had occurred, Gage was described as undergoing significant personality changes, such as a lack of inhibition in actions and speech, as well as profoundly poor and vulgar social judgment (Garcia-Molina, 2012). Gage's case was publicized by physician David Ferrier, who presented it as proof that the PFC governed functions associated with social and affective behavior. As would become evident in years to come, Ferrier was only partially correct.

The Early to Mid-20th Century

The early 20th century brought a greater understanding of brain function through the development of architectonics, or the differentiation of functional units of the brain using the types and functions of specific cells (Carlin, 2007). Through the study of cell structure, Brodmann developed his cortical map dividing the brain into discrete areas, which is still used today to standardize locations of function and injury. Architectonics also allowed for the discovery and mapping of the rich cortical connections between frontal regions and other areas of the brain (Benton, 2000). Another important occurrence was the outbreak of World War I and World War II, which offered tragic but unprecedented opportunities to study large samples of brain injured individuals over time. Researchers such as Feuchtwanger and Kleist created a wealth of data detailing case studies of the overall inability of patients with frontal injuries to regulate affect or to integrate and control behavior despite unimpaired cognitive functioning. The data collected from these patients led to the primary theoretical models of neurological function.

By the 1930's, prominent theories addressing the relationship between intelligence and behavior were being established. Raymond Cattell postulated his Investment Theory, delineating the difference between innate intelligence and acquired abilities. Cattell was influenced by the functionalist movement which asserted that behavior was the result of environmental influences, and deficits in ability were related to inequality in environmental factors more than innate genetic influences (Green, 2009). Cattell theorized intelligence consisted of both biologically based natural abilities and abilities acquired through positive and enriching environmental interactions. Using a new statistical procedure called factor analysis; Cattell developed a model separating intelligence into two factors; fluid and crystallized intelligence; the former reflecting novel problem solving and the latter reflecting acquired knowledge. Donald Hebb theorized a biological basis for reasoning skills by postulating a theory of how neurons aligned into optimal patterns of transmission in response to environmental stimuli (Bear, 2003). Hebb was describing what later came to be known as the concepts of neuroplasticity and long term potentiation. Hebb took the ability of the brain to reorganize after injury to what seemed a logical conclusion; if a portion of the brain was found to be dysfunctional, its removal would allow a more functional system to develop. Thus, Hebb became a leading proponent of the use of lobotomy to treat psychological and physiological disorders of the brain (Hebb & Penfield, 1940). Hebb supported his belief by citing that that psychometric assessment showed little changes in cognitive ability after frontal ablations. Ward Halstead was a psychologist who opposed this view because he observed first hand, wide ranging deficits in adaptive behavior in individuals with compromised prefrontal systems (Reitan, 1994). He was troubled by the inability of current assessment instruments to detect these deficits. Much like Cattell, Halstead used factor analysis to make a distinction between acquired skills or psychometric intelligence and innate skills based on the efficiency of biological processes. Biological intelligence, like fluid reasoning, comprised the innate ability of an individual to meet and overcome problems in everyday life and represented the overall efficiency of physiological processes within the brain (Reynolds, Castillo, & Horton, 2007). He reasoned that psychometric intelligence, defined as the performance of individuals on discrete psychometric tasks, was captured by current assessment methods but biological intelligence was not (Pallier, Roberts, & Stankov, 2000). Halstead empirically tested various tasks, assessing their sensitivity to brain injury to better ascertain an individual's true functioning. Through trial and error, Halstead developed an assessment battery based on a four factor theory of biological intelligence. The battery was later revised and researched by his colleague Ralph Reitan. Although the factors Halstead developed became anachronistic, the refined and well researched battery of tasks became the dominant paradigm of neuropsychological assessment in the United States (Reitan & Wolfson, 1993).

While Halstead was working to better understand functional impairments of individuals with brain injury, Soviet psychologist Alexander Luria, through his work with aphasic patients and brain-injured war veterans during World War II, was centered on developing a structured, integrated model of brain function (Luria, 2010). Luria believed higher mental processes evolved through highly complex functional systems that were fluid and changing with life experiences. Higher functions grew out of the simultaneous processing of cortical units and combinations of activation in these units differed based on the demand. A key point to this argument is that the patterns of interaction and the areas of involvement in the demonstration of a skill can differ based on age, culture, or life experience. Luria posited that the brain was arranged into functional units with more complex processes evolving as structural extensions of more primitive systems in response to environmental demands (Luria, 1973). In his study of brain injury, Luria concluded that injuries to specific areas at different times in development would result in highly different symptoms. Likewise, different abilities associated with more primitive cortices would either exacerbate or mitigate the effects of damage to higher cortices. Luria summarized the rationale for his beliefs by dividing the brain into three functional units: the primary unit, which regulated basic functions such as arousal and attention; the secondary unit, which received and organized sensory information; and the tertiary unit, which acted to plan, organize, and execute behaviors (Languis & Miller, 1992). The structure Luria imposed on neurological function allowed for a more cohesive understanding of the regulatory role of the PFC which he described as key to the difficulty of individuals with frontal lesions in initiating goal directed behavior and inhibiting a conditioned response (Banich & Compton, 2011). Luria emphasized the importance of the secondary unit in processing information and posited two methods by which it organized information; simultaneous and successive processing (Reynolds et al., 2007). Simultaneous processing refers to the synthesis of separate elements into groups and is often tested using spatial tasks. Successive or sequential processing involves maintaining a temporal order to information such as involved in a sequence of steps or rote memorization. Luria did not believe in a standard battery approach because each case was unique and required an individualized approach to assessment (Ardila, 1992). This was one of his three tenets of evaluation. Luria also believed that psychological processes had to be viewed as functional systems and not as discrete skills. He eschewed treating the PFC as a homunculus exerting a top-down only organization of lower functions. Instead, he saw all three functional units as interacting in systems where the efficiency of lower units contributed just as much as the PFC to the effective demonstration of a skill. He reasoned that damage or dysfunction in lower units therefore might mimic deficits in executive control. This led to the third tenant: that outcomes and quantitative scores were far less important than analysis of strategies and errors in task performance. A well-trained examiner could therefore intuit which functional impairment existed by examining the process by which an individual completed a task. In the 1960's, a group of practitioners in Boston studying patients with dementia and aphasias at the Boston Veteran's Administration Hospital were taking a very similar approach to assessment (Milberg, Hebben, & Kaplan, 2009). The group was led by Norman Geschwind, who had been studying aphasia from the neurological perspective. His largest contribution to neuropsychological assessment was the delineation of different aphasias though careful observation patients' task performance (Stringer & Cooley, 2002). These observations revealed qualitative differences in speech impairment not captured by quantitative outcomes and his work was key to validating the process approach to assessment. Edith Kaplan took her experience in this setting to develop the Boston process approach (BPA) to neuropsychological assessment (White & Rose, 1997). The BPA recognized the importance of capturing qualitative observations, testing the limits by giving additional cues and by giving options for untimed tasks (Milberg et al, 2009). Edith Kaplan was of the opinion that most standard and commonly used tasks could be used to assess neuropsychological functioning. For example, by analyzing the performance of an individual on a common assessment task such as the Wechsler Block Design test, one could simultaneously assess perception skills of Luria's first unit, integration skills of the second unit and planning and reasoning skills of the third unit (E. Kaplan, personal communication, February 22, 1996).

Influences on Modern Assessment Paradigms

The influence of clinicians such as Halstead, Luria, and Kaplan can be seen in many modern assessment instruments. The Halstead-Reitan Neuropsychological Battery, developed by Halstead and his student Ralph Reitan, is a fixed battery composed of tasks of reasoning, sensory, and motor skills that have been empirically proven to be sensitive to specific brain injury (Reitan &Wolfson, 1993). It became the dominant paradigm for neuropsychological assessment in the United States and is still used extensively today. The tasks were adapted for use in children ages 9 to 14 (Halstead Neuropsychological Test Battery for Older Children) and ages 5 to 8 (Halstead Indiana Neuropsychological Test Battery). Some of the tasks, such as the Trail Making task and Category test, have been adapted for individual use and in other neuropsychological batteries.

Cattell's student John Horn expanded Cattell's model to include an overall factor of general intelligence (McGrew, 2009). In the 1980's, John Carroll applied factor analytic statistics to validate further factors of intelligence that evolved into the Cattell-Horn-Carroll (CHC) model of cognitive functioning. CHC has become the dominant assessment model today, and, despite numerous revisions, still postulates a fundamental distinction between fluid and crystallized intelligence. Some current theoretical models recognize that Carroll's original description of fluid reasoning essentially equates it with current conceptualizations of executive functioning abilities (Miller, 2013). There has been much interest in incorporating EF and CHC theory into a cohesive theoretical model and there are indications that future research will result in a convergence of CHC and executive theory (Schneider & McGrew, 2012).

Edith Kaplan recognized the need to interface cognitive and executive assessment years before CHC became the dominant paradigm. She met skepticism over the use of qualitative data with the quantitative analysis and standardization of common qualitative observations in the BPA model. She assisted in adapting common cognitive batteries such as the Wechsler Scales to include supplemental norms and materials enabling them to be used as neuropsychological instruments that tested multiple levels of Luria's functional systems. She also developed other tasks such as clock drawing and word list learning tasks that included quantified scores for key qualitative observations thought to be indicators of brain injury. The Lurian and BPA emphasis on the process of task completion are reflected in test batteries such as the NEPSY-II, a flexible test battery with process oriented scores designed to assess children and adolescents in a variety of neuropsychological domains (Davis & Matthews, 2010), and the D-KEFS (Delis, Kaplan & Kramer, 2001b). The NEPSY II allows for differentiation between EF deficits and deficits in component neurological processes in populations six years and up (Homack et al., 2005).

Luria's model of functional units has also influenced the development of common cognitive assessment instruments such as the Kaufman Assessment Battery for Children, Second Edition (KABC-2; Kaufman & Kaufman, 2004). Naglieri engineered his planning, attention, sequential, and simultaneous processing model (PASS) based on Luria's functional units. His Cognitive Assessment System (CAS) is a cognitive ability test that minimizes cultural and language differences and maximizes measurement of information processing (Naglieri & Rojahn, 2004). Although the CAS based tasks on Luria's model of brain function and included some qualitative task analysis, they did not attempt to provide a basis for determining the frequency or pathology of such observational data. In the 1970's, the Luria Nebraska Neuropsychological Battery was created to compete with the dominant clinical assessment paradigm of the Halstead-Reitan battery. The Luria Nebraska tasks are composed of items that appear heterogeneous but that are designed to assess functional systems through quantitative and qualitative data (Purisch, 2001). Like the Halstead-Reitan, the tasks from the Luria Nebraska were extended down for use in children.

Relevant Measures of Executive Processes

The works of Luria, Cattell, Halstead, Kaplan and others have created a rich history of empirical validity for various measures of EF. To a large extent, most of the tasks used in this research study are based upon variations of categories of tasks that originate with these major influences in neuropsychological assessment. The most commonly used neuropsychological assessment instruments have incorporated variations on certain core tasks used for more than half century. A review of the general categories of these common assessment tasks is warranted.

Trail making tasks. A Trail making test is still part of the Halstead-Reitain battery and was originally used in assessment of United States army personnel in 1944 (Mitrushina et al., 2005). The individual is asked to draw lines connecting dots labeled with numbers and letters. Only the numbered dots are connected and must be in proper sequential order. On a subsequent trial, the individual is asked to alternate in sequential order between connecting dots with numbers and letters. The task is usually scored on time of completion but normative data for errors is available for some versions. The task is sensitive to visual perception, spatial skills, and speed of processing as well as to executive skills of spatial organization, problem solving and attention. Excessive errors have been correlated to dysfunction in the dlPFC and indicate a lack of executive inhibition. Other research has found it to be more aligned with the ability to switch between sets, a facet of cognitive flexibility, than with the ability to inhibit a conditioned response (Kortte, Horner, Carolina, & Windham, 2002). A version of trail making appears in the D-KEFS that contains five trials; a simple target cancellation task, sequential connection of numbers, sequential connection of letters, alternating between numbers and letters, and a simple motor processing speed task (Delis et al., 2001a).

Cancellation tasks. Cancellation tasks measure selective and sustained attention depending on the modality (Mitrushina et al., 2005). Sustained attention is measured in tasks requiring the identification of a reoccurring single target amongst distracter targets, while selective attention is measured through the matching of two targets in rows of distracter targets- for instance, finding two matching numbers in a row of numbers. These tasks are sensitive to anterior right frontal lesions as well as declines in processing speed caused by concussions. The WJIII COG NU includes a visual matching task with numbers, a visual matching task with pictures, and a target cancellation task (Woodcock et al., 2007) used to measure the broad CHC factor of processing speed.

Verbal and design fluency tasks. Verbal fluency tasks began to appear in formal test batteries in the 1960's and require rapid, timed recall of words based on a category or letter cue (Mitrushina et al., 2005). The task requires rapid development of an effective strategy, monitoring of performance and working memory skills. Verbal fluency tasks also involve temporal and parietal regions as shown by imaging studies and are heavily influenced by cultural and acquired knowledge. In the 1970's, design fluency tasks were developed as a non-verbal alternative. The task requires drawing novel designs using four lines. The lines must connect with each other on arrays composed of solid dots. Different versions of the task differ on the arrangement of dots and the presence or absence of non-target white or "empty" dots among the solid dots. This task has been shown to be sensitive to frontal lobe damage in the number of design and preservative errors committed. The D-KEFS includes both verbal fluency based on letter and semantic cues and design fluency tasks (Delis et al., 2001a). In addition, these tasks have switching component trials that involve switching between semantic categories for verbal fluency and switching from solid to white dots in design fluency. The tasks have been shown to be sensitive to frontal lesions with left frontal injury associated with deficits in verbal fluency, right frontal injury associated with deficits in design fluency, and general deficits in switching ability for both types of injury (Baldo, Shimamura, Delis, Kramer , & Kaplan, 2001). The NEPSY-II also includes a version of a design fluency task that does not require a switching component, and a verbal fluency task based on a categorical cue (Korkman, Kirk & Kemp, 2007b).

Category tasks. Category tasks were first incorporated into assessment batteries by Ward Halstead and requires the sorting of items based on different traits such as color and shape (Mitrushina et al., 2005). A chime or buzzer is sounded in response to a correct or incorrect sort. This task requires attention, visual-spatial detail orientation, and general reasoning ability most broadly associated with executive functioning. The Wisconsin Card Sort Test (WCST), developed in the 1940's, is another version of a category test where the participant must recognize a correct sort pattern has changed when the examiner begins to provide negative feedback to previously correct sorting categories. The hallmark of these tests is the process of active learning or the ability to modify one's cognitive set based on feedback (Miller, 2013). The WCST is sensitive to lesions in the dlPFC but performance can also be impaired with damage in the left parietal area accounting for the perception and integration skills necessary to complete this task successfully (Stuss & Levine, 2002). The WJIII COG NU includes a test of concept formation with examiner feedback that requires recognition of why certain shapes of differing size and color are grouped together while others are excluded (K.S. McGrew, Schrank, & Woodcock, 2007). The D-KEFS and the NEPSY-II both contain categorical card sorting tasks that do not provide examiner feedback (Delis et al., 2001a; Korkman et al., 2007b). The D-KEFS also includes a score for the accuracy of the verbal sort description.

Stroop task. John Ridley Stroop developed the Stroop task in 1935 to explore inhibition of over-learned responses, a concept first described by John Cattell and William Wundt (Mitrushina et al., 2005). It involves the rapid, timed naming of colors and words followed by a trial where the participant is asked to name the color of ink printed in incongruous color names (e.g. the word green printed in red ink). The task measures how well an individual inhibits the over-learned reading response. The task is sensitive to left frontal lobe damage and imaging scans have shown the expected involvement of both the ACC and PFC. The Stroop task was adapted for use in various batteries including the D-KEFS which added a third condition requiring switching between reading a word and naming the color ink based on a visual cue (Delis et al., 2001a). The NEPSY-II also includes a Stroop task with both inhibition and switching trials that uses shapes and arrows in lieu of words and colored ink (Korkman et al., 2007b).

Models of Executive Function

Based on the work of early neuropsychologists and cognitive theorists, various models have been developed to explain the intricacies of executive processes. The following is a discussion of the major emergent models that have influenced assessment practices.

The Supervisory Activation System (SAS)

In 1974, Baddeley and Hitch posited a three component model of memory consisting of a phonological loop, a visual-spatial sketchpad and a central executive that maintained information in short term memory for applying problem solving skills towards a perceived goal (Baddeley, 2000). To solve problems related to the interplay of verbal and visual information, the Baddeley-Hitch model was modified to include a working memory component called the episodic buffer, which represented a temporary storage unit that allowed for multi-dimensional coding of information (Repovs & Baddeley, 2006). Later, the central executive governing conscious problem solving was posited as a supervisory activation system (SAS) for these processes. Its function was described as overcoming learned, automatic responses by directing conscious attention to a situation not responding to current behavioral routines (Norman & Shallice, 1986). The episodic buffer is a function of the central executive that, along with a supervisory attention system, holds mental set information relevant to goal orientation while resisting distractions (Coolidge & Wynn, 2005). This allows multiple options to be considered simultaneously against a stable goal construct, which supports temporal stability of set, a critical component for fluid reasoning. Fluid reasoning therefore may be defined in terms of a process whereby working memory, directed by an SAS, provides guidance towards a goal. When goal attainment is realized, cortical systems governing behavioral reinforcement trigger emotional salience. The Baddeley and Hitch model and the Norman and Shallice SAS were originally conceptualized to reflect a unitary model of executive functioning. In subsequent research and clarification on a supervisory executive system, both models indicate that evidence exists of localized processes such as attentional switching, cognitive flexibility (monitoring and checking), and top-down maintenance of schema in the face of distraction (Stuss & Knight, 2002). However, these skills are still considered part of a holistic system that is only activated when a task can no longer be addressed by default problem solving systems. Measurement of individual aspects of executive processes will vary with both the efficiency of perceptual and integrative skills, as well as with the specific task presented.

Lezak's Four Executive Processes

Muriel Lezak, a strong proponent of the flexible battery approach to assessment proposed four major classes of executive processes: formulating goals, planning, carrying out plans, and performing these activities in a meaningful, efficient way (Lezak, 1982). Formulating goals is most dependent on recognizing and interpreting the meaning of environmental cues, and possessing the motivation to respond by initiating behavior. Planning is dependent on sustained attention and the ability to form conceptual frameworks to guide activity. Carrying out plans requires maintenance/modification of one's approach through uncertainty, switching (avoidance of preservative actions), and inhibition (ability to stop an intended series of responses). Generally the more open ended a task, the more evident problems with carrying out plans will become. Finally, efficiency in these processes is governed by selective attention and self- monitoring; the ability to perceive one's errors and to act to correct them. An important difference in Lezak's model is that not all functions of the frontal cortex are considered executive in nature, and activation of the frontal cortex does not necessarily mean the task is sensitive to frontal impairments (Royall et al., 2002). Non-executive systems that involve frontal activation can help increase efficiency of the processing and organization of information. For example, although prefrontal areas such as the ventromedial cortex are activated in simple memory tasks requiring cued recall of information, these skills are rarely significantly affected by pre-frontal lesions. It is only when information must be sequenced, organized or maintained in the presence of interference that deficits become evident.

McCloskey's Assessment Model

McCloskey et al., (2006) define EFs in similar terms as Lezak. Executive processes must involve frontal activation but also must direct or cue other mental processes. Executive processes operate to regulate the four domains of perception, emotion, cognition, and actions and operate at different systems' levels such as intrapersonal, interpersonal, and environmental. The model proposes 23 distinguishable functions of executive processes that are evaluated based on a hierarchical processing approach. By making detailed qualitative observations of how cognitive tasks are engaged, assumptions can be made regarding executive processes. These assumptions can be further tested with executive assessment tasks and naturalistic observations in multiple settings.

Damasio's Somatic Marker Hypothesis

The use of terms such as *supervisory activation system* or *metacognition,* tends to focus on a dualist view of mind and body and treats the PFC as a homunculus. However, overall self- regulation and monitoring of behavior involving frontal activation is only partially based on conscious processing. Damasio (1994) explains this process by stating that each available existing response to a situation has a somatic marker that creates a visceral emotional response making it more or less likely to be inhibited. The Somatic Marker Hypothesis (SMH) postulates that the brain encodes the frequency and emotional salience of past knowledge and events; therefore, people will exhibit emotional bias in decision-making. They will tend to rely on simplified patterns of decision-making or heuristics and employ more automatic patterns of response than to use logic and analysis to solve problems. For instance, in estimating populations of countries, people tend to overestimate populations of countries with which they are familiar and underestimate populations of those they are not (Brown & Siegler, 1992). In such cases, decisions are quick and automated rather than involving executive processes for higher reasoning skills. To stress the importance of emotional salience in decision-making, Damasio used the case of Phineas Gage to illustrate the widespread impairment in social learning and judgment caused by damage to the vmPFC, and further research has shown that patients with bilateral damage to this area evidence profound impairments in decision-making and initiation of action (Bechara, Damasio, & Damasio, 2000). In the absence of pertinent information for novel problem solving, heuristical patterns of automated decision-making take over. This is best exemplified in research on how patterns of positive reinforcement enhance cognitive performance in the absence of explicit memory (Frank, O'Reilly & Curran, 2006). Participants in the study who were administered a learning task before and after being given an amnesiac drug were still able to implicitly identify correct responses in post assessment taught to them if positive and negative feedback had been provided in the baseline administration. The participants could not explicitly explain why the answers were correct, only that they "felt" right. Thus, correct response patterns were independent of conscious executive control.

Damasio's theory has widespread implications for the involvement of executive processes. First, there is the implication that overreliance on heuristical patterns of thinking would result in decreased self-monitoring and frequent errors in judgment. Second is the implication that activation of executive processes such as conscious response inhibition is reduced in situations with high emotional salience. This is one possible explanation for the differences in observation and direct assessment of executive functioning.

Hot and Cool Executive Processes

Building upon the anatomical separation of self-regulation in the PFC, some have attempted to delineate executive processes into categories depending on what facet of problem solving they govern (Bechara et al., 2000). Because of this, some models have chosen to distinguish processes associated with cognitive control, or "cool" executive functioning and those regulating behavioral and affective functioning or "hot" executive functioning (Hongwanishkul, Happaney, Lee, & Zelazo, 2005). Cool executive processes are conceptualized as those assessed through standard cognitive and neuropsychological assessment batteries. Hot executive processes are assessed through tasks that center upon delayed gratification and risk-reward association. However, this approach has met with only moderate success. The distinction seems to be most evident in younger children; however, facets of self-regulation described as "hot" EFs have not been shown to be consistently related to either academic competency or behavioral regulation over time (Brock et al., 2009; Hongwanishkul et al., 2005). Denham et al.

(2012) proposed a two-factor model consisting of executive compliance and impulse control, with hot and cool processes comprising sub factors within impulse control. This model supported teacher observations of pre-school behavior and successfully differentiated social competence in children; however, no distinction was made between the overall contribution of hot and cool processes. The BRIEF rating scales also make a distinction between cognitive and behavioral executive control, although further analysis has shown some re-alignment of the original factor scales in a clinical sample (Gioia et al., 2002).

The idea of hot and cool executive processes that can be evaluated separately is a useful paradigm that can readily explain differences in direct assessment and real world observations of executive processes. However, there seems to be little evidence that such a distinction can be used to predict cognitive and behavioral competencies. Damasio's theory and subsequent research seems to indicate the opposite; emotional and cognitive facets of self-regulation constantly interact in decision-making and cannot be evaluated as separate entities. This contention is supported by a meta-analysis of neuroimaging studies that indicated widespread hypoactivation in both emotional and cognitive centers of regulation in the PFC (Dickstein et al., 2006). Thus, impairments in a combination of hot and cool processes that impair functioning should also be detectible, at some level, in standardized formal assessment.

Models Germaine to the Current Research

As mentioned previously, numerous executive processes have been described in the research literature (Packwood et al., 2011). However, throughout the historical and conceptual discussion of EFs, certain constructs reappear with regularity. Concepts such as a supervisory system for conscious attention, inhibition of over-learned responses, and flexibility in solving novel problems by forming, executing and monitoring effective strategies are recurring themes. Differences in the conceptualization of whether these are separate or unitary constructs, how they should be measured and if other factors such as verbal versus spatial modalities of assessment, have been a major impetus to theoretical agreement. Given the instruments being evaluated and the theoretical frameworks in which they were developed, it is likely that certain factors common to previous research will emerge. Determining the most likely resultant factors requires analysis of three related paradigms in current theory.

CHC Theory

As mentioned previously, Raymond Cattell used factor analysis to separate the concept of intelligence into two factors based on his Investment theory (Zhang et al., 2012). His student, John Horn, expanded the model to three factors to capture an overall measurement of intelligence or the “g” factor. In the 1980’s, John Carroll applied factor analytic statistics to validate a theory of intelligence that evolved into the Cattell-Horn-Carroll (CHC) model of cognitive functioning. Carroll’s analysis expanded the model to three strata; stratum I represented an overall measure of “g”, Stratum II consisted of eight broad domains of cognitive function, and stratum III represented 70 narrow abilities. Using the Cattell-Horn model as a basis for his research, Carroll conducted analyses on over 460 historical datasets creating what John Horn described as the psychometric equivalent of the periodic table of elements (McGrew, 2009). The two theories merged to become the Cattell-Horn-Carroll (CHC) theory of intelligence, which has become the theoretical basis for current versions of many cognitive batteries including the WJIII COG NU, the K-ABC 2 and others. Currently, nine Stratum II factors are commonly assessed and revisions in both broad and narrow stratum abilities have been enacted and/or proposed.

Although CHC theory does not directly address assessment of EFs, it includes factors, such as short term/working memory (*Gsm*) and fluid reasoning (*Gf*), that have been theorized to be part of, or heavily mediated by, executive processes (Miller, 2013). For instance, the description of *Gf* provided by McGrew (2009) as "the use of controlled mental operations to solve novel problems…linked to cognitive complexity which can be defined as a greater use of a wide and diverse array of elementary cognitive processes" (p. 5) mirrors the concept of an executive SAS. Working memory or the manipulation of data in short term memory has also been described as a phenomenon that arises from the interplay of the CHC short-term memory factor (*Gsm*) and executive processes that inhibit interference and response competition (Coolidge &Wynn, 2005). The view that working memory is an emergent, higher construct more complex than short term memory, is supported by brain imaging studies that consistently show higher activation and involvement of multiple areas of the PFC during memory tasks where information must be manipulated versus those assessing rote memory (Carpenter et al., 2000). There is also evidence in research that working memory and executive tasks merge into a single construct when subjected to factorial analysis (McCabe, Roediger, Mcdaniel, Balota, & Hambrick, 2010). Similar studies have generally found the two factors to be discernible but highly correlated (Garlick & Sejnowski, 2006).

School Neuropsychological Processing Model (SNP)

Factor analytic studies examining the relationship between CHC factors and neuropsychological instruments have largely concluded that common executive assessments fall into theorized CHC factors and do not measure separate constructs (Floyd et al., 2010). However, this conclusion was somewhat simplistic given the results obtained. When cognitive tasks from the WJIII COG NU and executive assessment tasks from the D-KEFS were combined in factor analysis, a fluid reasoning construct failed to emerge. Instead, significant loadings and cross loadings occurred between factors representing executive functioning, processing speed, and working memory. Further, these domains were highly correlated with overall intelligence. An additional anomaly was that verbal and non-verbal tasks of reasoning from both instruments inexplicably loaded on a crystallized intelligence factor. One possible reason for the confounding results obtained is that each cognitive task required some degree of executive processes and variance due to these processes confounded factor loadings. By including all tasks from both batteries, the authors failed to take into consideration the concept of task impurity, which states any executive process can only be measured by its ability to modulate other non-executive processes (Jurado & Rosselli, 2007).

Including executive measures in analysis with a wide array of cognitive tasks known to primarily tap non-executive processes ignores the fundamental difference in the theoretical constructs. For instance, when the Woodcock Johnson achievement and cognitive batteries are merged together in an EFA without regard to their respective constructs, a markedly different factor structure emerges that makes it difficult to discern which tasks assess predominately cognitive skills versus academic skills (Dombrowski & Watkins, 2013). This is to be expected since every academic task taps multiple cognitive abilities just as every cognitive task taps some degree of executive processes.

A better conceptualization of how CHC theory and EF assessment intersect can be found in the SNP model (Miller, 2013). The SNP model uses Lurian architecture to categorize skills assessed into the areas of sensorimotor functioning, basic cognitive functions, facilitators of basic cognitive functions, and acquired knowledge. It involves a process-oriented approach to identifying strengths and weaknesses in these areas using CHC assessment theory as a basis. For instance, the SNP model views the CHC factor of fluid reasoning as interchangeable with the measurement of executive problem solving functions and defines narrow band abilities including concept formation and category switching. The CHC model also recognizes that skills such as attention, processing speed, and working memory represent intersections of a wide array of basic cognitive processes and the executive skills that govern them. In describing these processes as cognitive facilitators, the model recognizes that their effect on the efficiency of problem solving can be inferred through various methods including direct measurement, process oriented observations of cognitive tasks, observer rating scales and naturalistic observations. A trained professional can therefore infer at which level of processing poor task performance might be most affected.

Miyake and Friedman Model

One way of answering criticisms has been to conduct factor analytic studies with longitudinal data to analyze changes in executive functioning. Miyake et al., (2000) conducted such research in hopes of developing a parsimonious model of executive processes that reflects developmental changes over time. Their methodology was to identify tasks that assessed simple, discrete skills underlying more complex executive constructs such as concept formation and planning. Subsequent studies on twin populations (Friedman et al., 2008) using the same analytic methods have verified the original model which is now based on four general conclusions (Miyake et al., 2012). First, executive processes exhibit both unity and diversity. Three general factors, updating (monitoring and updating information in working memory), flexibility/switching, and inhibition, were identified. However, when a general second order factor of executive functioning was identified, inhibition no longer explained separate variance in the model. Second, the unitary factor of EF has shown validity in predicting both cognitive and social aspects of self-regulation. Third, executive processes, especially the unitary latent variable, show high hereditability. Fourth, executive processes show stability over time with longitudinal data revealing moderate to high correlations in self-regulative ability from early childhood to adolescence. The three factor unity and diversity model has proven difficult to replicate in younger children where individual factors have not been shown to be as discernible as in adolescent population (Lee et al., 2013).

The Miyake and Friedman model is important because it highlights three core skills in the regulation of cognitive processes: inhibition; which correlates to a unitary executive construct; set-shifting, which underlies flexibility in problem solving; and register updating, which involves effective gating and controlled retrieval of information from long-term memory. It also highlights a developmental progression and increasing delineation of these skills with development. In addition, it exhibits an expected inverse relationship between switching and inhibition that shows the juxtaposition of these skills in higher reasoning and problem solving (Miyake et al., 2012). The model has also been shown to reveal the contribution of these skills to more global measures of executive functioning such as the WCST, and to applied skills such as mathematics ability (Bull & Scherif, 2001). It remains to be seen if the assessment tasks used in this research assess these proposed skills in a similar way.

Neurology of Executive Processes

It is apparent that involvement of the PFC in problem solving is not “all or nothing” but rather occurs on a continuum that depends on both task complexity and novelty. To understand how and when these skills are activated requires an examination of the major neurological systems that act to govern behavior and cognition. This examination begins in the limbic system, or where the brain's separate perception and action circuits first interface with one another to direct adaptive behaviors (Goldberg, 2001). Key structures in learning and emotion include the amygdala, involved in approach/avoidant response, specifically in reading the emotional content of facial features; and the hippocampus, which is crucial to the formation and recall of memories (Garrett, 2009). Components of the limbic system govern emotional response to stimuli and interface with the structures of the basal ganglia (BG), which connect to the motor cortex to initiate or inhibit programmed motor movement.

The PFC interfaces with limbic areas through direct connections and through the caudate nucleus and globus pallidus, which are structures of the BG that play important roles in processing positive and negative feedback to behavior (Faw, 2003). The critical role of the pathways of the caudate nucleus and globus pallidus to the interface of emotions and motor control is clearly evidenced in Huntington's disease where damage to these systems results in both severe apathy and complete, often fatal, paralysis (Kailasanath & Fu, 2010). The complexity of connectivity between the PFC and other cortical units limits this paper to the discussion of the major cortical systems involved in voluntary and automated behavioral and cognitive processes. Although sensory and associative functions of perception and information integration are crucial to every functioning, the complexity of those systems is beyond the scope of this literature review.

It is more pertinent to understand the integration of the major executive systems and to understand how cognitive processing both initiates and inhibits behavior.

The Motor and Subcortical Systems

The motor and oculomotor circuits govern supplementary/primary motor movements and visuospatial and visual-perceptual functions respectively. The motor circuit seriates and organizes motor movement and connects to the globus pallidus and caudate through the putamen (DeLong & Wichmann, 2010). The oculomotor circuit originates in the frontal eye field and frontal and posterior areas of the parietal lobe and also has connections to discrete areas of the globus pallidus, caudate nucleus and thalamus. Deficits in these lower functions can result in deficits in motor control and visual attention/scanning respectively.

The motor circuit is composed of primary and supplementary motor areas with the supplementary motor areas further divided into the lateral premotor area and the medial supplementary area. The primary motor cortex organizes incoming sensory information from associative areas of the lower cortices and executes motor movements. Premotor and supplementary motor areas can be distinguished by their influence on the motor cortex: the premotor area is more involved in novel motor sequences in response to the environment, while the supplementary area governs more automated motor sequences.

The motor and occulomotor cortices do not share a direct connection to the systems of the PFC. It is important to note that all major frontal-subcortical loops are closed loops that do not share connections with each other. At first, this may seem confusing since these entire systems act in concert to govern and execute cognition and behavior. However, all of these systems have common connections to structures of the limbic system, basal ganglia, thalamic regions and associative cortices. Although each is a discrete system, they influence thought and action through their activation or inhibition of subcortical systems. For instance, the dlPFC connects to structures of the basal ganglia and loops to the primary motor cortex through subcortical loops to the ventral anterior nucleus of the thalamus. Therefore, it exerts influence on the cortical motor loop but by no means directly connects to or controls it (Faw, 2003).

In discussing the major executive loops of the frontal cortex, it is important to remember that these systems act more as a committee, each having a weighted “vote” that determines whether action and thought is inhibited or initiated. This indicates that even though the PFC governs thoughts and action, the resultant behavior is solely dependent on activation in subcortical loops. The term “circuit” highlights the fact that observed executive dysfunction might have nothing to do with impairment of the PFC but rather arise through dysfunction in subcortical systems. As an example, abnormal impulsivity, euphoria, and impulsivity common in damage to the OFC can also arise from lesions in caudate and thalamic regions (Thimble, 1990).

The Anterior Cingulate Circuit: The “Pre-executive” System

Directly posterior to the frontal cortex and intricately tied to its function is the ACC. The ACC’s connections to the dlPFC, occulomotor centers, and the reticular arousal system make it a primary center for communication between the PFC and the midbrain (Paus, 2001). The ACC connects to the globus pallidus through the nucleus accumbens, a key structure for reward/consequence associations and the ability to delay gratification (Faw, 2003). The level of task difficulty seems to be the primary predictor for activation in the ACC, which may detect response conflict (when a given response is being unsuccessful) and raise the level of attention and stress. Thus, it may be the ACC that first detects failure of automated strategies and initiates conditions necessary for the PFC to engage conscious reasoning processes to evaluate, inhibit and/or alter pre-existing environmental responses (De Pisapia et al., 2007). This is consistent with the theoretical view of a central executive network which acts to inhibit the default network proposed by Mesulam (2002). Decreased activation of areas of the vmPFC and the posterior cingulated cortex are seen in tasks requiring a higher degree of problem solving and greater load on working memory. This is viewed as an inhibition of default routines in favor of a higher degree of involvement from the dlPFC (Greicius, Krasnow, Reiss, & Menon, 2003)

Emotional Executive Control: Orbitofrontal/Ventromedial Circuits

Intricately tied to the limbic systems and ACC, the OFC governs the conscious awareness of feelings, expression of empathy, and internal emotional states through connections to the hypothalamus (Cummings & Miller, 2007). Damage to the OFC often results in an inability to process the emotional content of situations resulting in socially inappropriate behavior. The OFC is the primary area involved in the somatic marker hypothesis, which states that bio-regulatory processes facilitate decision-making by assigning emotional “weights” in the form of visceral responses to options in making a decision (Damasio, 1995). More recently, a distinction has been made between functions of the OFC and of the vmPFC in emotional regulation. OFC neuronal activation coincides with emotional valuation of external stimuli while vmPFC activation responds to intrinsic subjective judgments of reward value including levels of satiation or desire (Bouret & Richmond, 2010). This indicates a complex interplay between identifying external factors salient to reinforcement and internal attributions, beliefs and desires in initiation or inhibition of motor responses. This distinction is further supported by research showing specific damage to the OFC is associated with increased aggression due in part to unpredictable reward/consequence contingencies (Leonard-Zabel & Feifer, 2009).

Cognitive Executive Control: Dorsolateral Prefrontal Cortex

The dlPFC mediates attention and focus, and regulates the shifting of set when flexibility in reasoning processes is necessary (Powell & Voeller, 2004). The dlPFC is the center of complex reasoning with connections to temporal and occipital sensory and perceptual centers in the brain (Cardoso-Leite & Gorea, 2010). A critical function of the dlPFC is to maintain a representation of a goal state independent of immediate stimuli. This allows for comparisons and analysis of how possible strategies might affect goal attainment (working memory) as well as monitoring of competing response patterns (task switching). Primate studies have shown that the greater the demand for comparisons involving discrimination and sorting of stimuli, the more activation switches from the midbrain to the dlPFC (Antzoulatos & Miller, 2011). Thus, task difficulty and complexity results in greater dlPFC activation and less reliance on pre-programmed responses. A key difference in the neurons found in the dlPFC is their ability to maintain information in the presence of distracting stimuli whereas neurons in the inferior temporal regions cease firing when distracted by a new stimulus (Arnsten & Li, 2005). The mutual interaction between the ACC and dlPFC can be viewed as a critical link in managing and evaluating the interplay between automated responses and increased executive involvement in tasks of fluid reasoning.

Inferior Frontal Cortex and the "Kill Switch"

The inferior frontal cortex (IFC) is located medial to the OFC and has a role in inhibition of in-process actions (Aron, 2008; Chambers, Garavan, & Bellgrove, 2009). The function of the (IFC) – sensory motor cortex - subthalamic circuit works by activating a "kill switch" that involves a hyper-direct path to the thalamus while suppressing activation of structures of the basal ganglia that are responsible for disinhibiting a planned motor response. Support for this function can be found in primate studies where damage to the IFC results in deficits in learning primarily due to increased preservative errors (Kowalska, Bachevalier, & Mishkin, 1991). Components of the left IFC which include Broca's area, have been implicated in reading deficits and seem to perform a discriminatory and inhibiting function in converting orthographic information to representative sounds (Burton, LoCasto, Krebs-Noble, & Gullapalli, 2005).

The Role of Long Term Potentiation (LTP)

As outlined in the previous section, executive processes are involved to differing degrees in the regulation and inhibition of automated processes. The strength of these automated, habitual responses is an important factor in when and to what degree EF promotes self-regulation.

Habitual environmental responses develop because of a process known as long-term potentiation (LTP). LTP is a process where dendritic connections are thickened and made more sensitive to the neurotransmitter glutamate (Rattan, 2009). This process dramatically increases the likelihood of neurotransmission along a specific pathway. LTP has long been known to be important to the formation of episodic and semantic memories in the hippocampal region (Faw, 2003); however, awareness of how this process maintains behavioral routines making them more difficult to consciously inhibit is a fairly recent development. LTP is now known to be involved in the strengthening behavioral routines and emotional responses in the basal ganglia involving reward and gratification responses (Berretta, Nisticò, Bernardi, & Mercuri, 2008). The more habitual these routines and the stronger the neural connections, the more resistant they are to being overridden or inhibited by executive processes. In other words, the more habitual the response, the more environmentally dependent it is. This is further evidence that execution of behaviors is dependent on lower cortical systems and one possible explanation why differences in observer ratings and assessment of executive competency are observed. Behavioral dysregulation observed in "real-world" processes may be less a function of specific EF deficits and more a function of reinforced maladaptive behavioral routines that "override" executive influence.

Child Development

To properly understand the need for differentiation of child and adult models of EF, one must understand the remarkable neurological growth and changes that take place from gestation to adolescence. From ages 3 to 6, children experience marked spurts of PFC dendritic growth and cortical volume, with myelination of these connections in secondary areas peaking by age 7, and in the PFC during adolescence. In addition to growth, there are also periods of thinning and pruning of PFC to mid-brain connections (Spencer-Smith & Anderson, 2009).

Cognitive skills, such as working memory, or processing speed, do not always develop on a linear timetable. Selective attention skills show rapid development from age 3 to 6 with another rapid improvement between 8 and 10 years of age. Infants as young as 14 months evidence self-regulation skills, which also show rapid development in early childhood that continues until pre-adolescence. During development, the process of synaptogenisis and subsequent pruning make pathways between the PFC and other parts of the brain more efficient, and the repertoire of existing responses becomes greater as LTP creates patterns of implicit and explicit learning. The more experience available, the more so-called "novel" problems will trigger familiar cognitive and behavioral patterns which can be "adjusted" through executive processes to facilitate new strategy development.

Maturation of executive functioning, in concert with increases in automatic processing speed, explains a great deal of variance in task performance as children age. For instance, older children suffer less of a response cost as tasks become more complex than younger children do (Davidson, Amso, Anderson, & Diamond, 2006). Also, after age 12, increased accuracy in matching tasks has been found to cease correlating to decreased completion time (Best, Miller, & Naglieri, 2011). These results represent both an increased ability to develop effective task strategies as well as more efficient basic processes requiring less executive monitoring.

Neurologically, these results mirror increased myelination and cortical pruning allowing for more efficient and accurate task performance. An important question is how to best summarize these profound neurological changes in terms of the development and definition of executive processes. The first step is to emphasize that development is a fluid but predictable process in which latent skills are not seen as cognitive leaps but rather as ever present skills that become more clearly definable with development. In other words, stage theories of child development may be a good way to summarize skill acquisition but may mask the fluidity of true developmental trajectories.

Connectionist Theory and Emergence of Executive Processes

Historically, developmental theories have been defined in terms of stages. Piaget developed his well-known stages of cognitive development based upon assumptions that transitions to qualitatively different levels of processing occurs in stages and that these changes are abrupt, simultaneous across domains, and occur in a fixed progression

(Mareschal & Schultz, 1999). However, although research supports most of Piaget's demarcation of abilities, there is evidence that these skills do not suddenly emerge but are evident much earlier in less developed form. For instance, in Piaget's demonstration of seriation development, children were given sticks of different lengths to order by size (Inhelder & Piaget, 1969 in Mareschal & Schultz, 1999). Piaget found that children first began to order sticks by size at 5 years of age but did not combine them to form a single pattern. By the age of 7, accurate seriation skills "emerged" in children.

Subsequent research has shown that very young children demonstrate precursor sorting skills, especially when the size between items is expanded (Mareschel & Schultz, 1999). Seriation, therefore, is a latent and developing skill that develops incrementally until it is readily observable. This and other research showing the early presence of underdeveloped skills have cast doubt on the stage view of development in favor of more fluid connectionist theories. Connectionist models are based on the premise that there are innate patterns of neurological connections between processing units within the brain (Munakata & McClelland, 2003). These connections are "weighted" in that some connections are more strongly associated than others are. As learning occurs, adjustments are made to these weighted values of these connections through LTP in observational learning and through its opposing process of long-term suppression (LTS) through learning from error.

Connectionist models have the advantage of using computers to build simple neural nets and then to present these neural nets with specific problems to see how the connections change to develop and employ a solution. Through the use of these networks, possible underlying neural representations of object permanence help to explain differences in infants' performance under variant task conditions. For example, a difficulty with explanations of object permanence in stage theory has been that although infants will not reach for items that have been occluded from sight, they will reach for items to which they have been exposed that have been occluded in darkness (Munataka, Mclelland, Johnson, & Siegler, 1997). Neural network simulation has revealed a likely explanation for this phenomenon by positing that the strength of object representation in the temporal cortical area is present much earlier than Piaget anticipated thus giving the child the ability to form and maintain a trace memory of an object that can no longer be seen in the dark. However, although the child's working memory neural network is developed enough to sustain representation of the object occluded in darkness, it is not developed enough to sustain the image when the visual interference of an occluding object is presented. Viewing object permanence as a developmental function of working memory and goal maintenance is more closely aligned to current understanding of executive processes. In addition, it reveals that, although executive tasks may not reveal their presence, these processes are present and central to governing learning and self-regulation in the earliest stages of development.

Connectionist models also explain the interaction between the processes of automated responses governed by the ACC, and the supervisory executive system employed when automated processes are unsuccessful (Ruh, Cooper, & Mareschal, 2005).

Rather than an "all or nothing" model of executive involvement, there are degrees of executive involvement depending on the complexity of the problem as well as in the efficiency of lower cognitive functions. The dynamic interaction between these systems explains how simple routines can be changed, altered, or abandoned by periodic attenuation of the PFC to help maintain goal orientation in sequences of routine tasks. Connectionist models describe child development in terms of a dynamic interaction between biology and environment that creates the most adaptive processes for the survival of the person. This indicates that although differences in ability levels should be present across children due to environmental factors, similarities in the progression of development of executive processes should both be present and reflecting of observed changes in cognitive and behavioral abilities. This contention is supported by three key areas of research: neuropsychological assessment, neuroimaging, and studies showing the epigenetic effects of environment on cognitive development.

Evidence for Executive Constructs Across Development

Among the executive constructs presented, general observation of toddlers and pre-school children leaves little doubt that inhibition is the first self-regulatory process to emerge. In formal assessment of preschool children, research with differing instrumentation and methodology has found inhibition to be a central factor in executive processes. The ability to self-regulate or inhibit behavior in formal assessment has been found to show dramatic gains between 3 and 4 years of age with almost a standard deviation change in ability (Willoughby, Wirth, & Blair, 2012). Furthermore, initial rates of measured executive processes showed constant rates of change from 3 to 5 years of age. Not surprisingly, assessment was more effective at revealing deficits in self-regulation than it was in differentiating average to above average skills. The fact that changes were stable over time for individual participants supports the contention that hereditability plays an important role in development of executive processes.

In a another study of 3 to 5 year old children, inhibition and working memory were distinguishable as separate constructs but were highly correlated (Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012). The authors summarized their findings in light of similar research and concluded that differentiation of executive processes begins to occur in the early preschool years with a unitary executive structure explaining less variance in performance with increased age. The process of cortical thinning and pruning which takes place during this early developmental period was found to correlate with performance on some, but not all, measures of executive processes in children 8 to 19 years of age (Tamnes et al., 2010). Specifically, tasks associated with working memory and response inhibition were associated with thinner cortex areas bilaterally in frontal and posterior brain regions, emphasizing that concepts such as inhibition and working memory represent systems of interactions between frontal and posterior regions.

The differentiation of executive processes also mirrors developmental patterns of cognition. In younger children, inhibition has been identified as a stronger determinant of problem solving ability as opposed to working memory as the primary predictor in older children (Senn, Espy, & Kaufmann, 2004). A key aspect of problem solving, shifting, or cognitive flexibility, has been difficult to define as a construct in childhood studies (van IJzendoorn, Bakermans-Kranenburg, & Ebstein, 2011). Theoretically, both inhibition and set shifting are foundational to problem solving in that they allow for resolution of competing responses based on a stimulus. However, whereas inhibition requires simple suppression of an overlearned response, set shifting requires a choice between competing responses. In most assessments, set shifting is assessed by having the participant assess and choose between conflicting responses over time. Thus, processes such as set shifting, flexibility and inductive reasoning require increased attentional control and flexibility in concert with inhibition and working memory skills. There is evidence that attentional control does not undergo appreciable linear change until early to mid-adolescence (Spencer-Smith & Anderson, 2009). This explains why children seem to exhibit more plasticity than adults after a TBI. However, the truth may be that the more profound effects on attentional shifting and flexibility from injury may not be apparent until adolescence.

Overall, strong evidence exists in assessment research for a consistent developmental pattern of increasingly differentiated executive systems in children consistent with connectionist theory. Differences in the development of executive processes, and the development of habitual reactions that resist executive control, are best explained by environmental factors.

Epigenetics: Building on Connections in the Brain

All children undergo a similar pattern of skill development exemplifying the innate development of cognitive systems. Within a range of time, a majority of normal developing children will utter a meaningful word in relation to an object (12 months), place objects into simple categories (15 months), and by 20 months should develop visual-spatial discrimination allowing for construction of objects (e.g. building with blocks) and relative distance (e.g. "here" versus "there" (Linder, 2008). These skills emerge with environmental stimulation and can be delayed by biological and environmental impairments. Many differences in skills not due to genetic defects can be explained by bio-psychosocial interaction effects or epigenetics.

Epigenetics refers to the bio-chemical processes that regulate the phenotypical expression of genes within cells to support differentiated functioning (van IJzendoorn et al., 2011). Methylation is a primary mechanism for epigenetic changes and involves the attaching of a methyl molecule to cytosine, one of the four main bases in DNA, which inhibits transcription factors that promote trait expression from access to the gene. The effects of methylation are copied each time the gene is reproduced. The process of methylation is important to the correct expression of cell function as the same biochemical processes support differing functions within differing organs.

According to Hochberg (2011), the expression of individual genes and therefore alterations in phenotypical expression occur within an individual life span as a result of both environmental cues and immediate adaptive responses (IAR). IAR occur as a result of extreme deprivation or environmental stress and operate to preserve survival whatever the long term cost to the individual. For instance, a developing fetus will alter phenotypical expression in the face of maternal malnutrition to devote more energy in the development of vital organs and less to secondary systems such as higher cortical units. A good example of the extreme effects of IAR can be found in the survivors of the Dutch famine of World War II. Children exposed to famine for a period as short as six months were suffered a significantly higher risk for physical illnesses, such as obesity and glucose intolerance; as well as for affective disorders, such as schizophrenia and depression. A major phenotypical response to the environment involves the stress mechanism. Stress is meant to signal the lack of homeostasis with the environment and acts to promote remediation. The body's stress reaction involves stimulation of the amygdala, suppression of hippocampal activity resulting in loss of sleep and hyper-arousal. Conditions, under which stress becomes acute, especially in developmental phases of brain growth, can significantly, if not permanently impair cognitive functioning, emotional regulation and hypersensitivity to somatic feedback from the body.

The study of epigenetics arguably can be construed as explaining a great deal of variance in cognitive and executive processes between individuals. As mentioned previously, the glutocortoid expression gene affects regulation of chronic stress levels and anxiety. Rat studies have shown that females with increased licking and grooming attention to offspring have increased levels of the inhibitory neurotransmitter, GABA, resulting in increased regulation of the emotional response in the amygdala (Hackman et al., 2010). Further, this increased nurturance results in increased hippocampal activation and increased biological precursors to future synaptogenesis. Thus, in situations of high parental nurturance, both immediate and future memory and learning processes are enhanced. In rats with low maternal care and high stress pregnancies, the opposite developmental pattern is evident. The processes are more resistant but not immune to sudden environmental changes. For example, evidence exists that sudden negative environmental factors such as physical or sexual abuse in children can result in severe depression caused by chronic patterns of memory disturbance due to hippocampal methylation.

Biological genetic differences have also revealed some people are more susceptible to environmental stressors than others are. Such differences are exemplified by variation in the alleles on a gene that regulates the production of 5-HTTP, the transporter protein for serotonin (Meaney, 2010). People with short alleles show increased risk to depression, anxiety, impulsivity, and heightened fear response; however, a nurturing, supportive family environment has been found to counteract the effects of this genetic anomaly. In addition to gaining insight into the biopsychosocial development of depression, anxiety, and aggression, epigenetic studies have sought to explain disabilities such as autism in relation to links between various environmental factors including exposure to heavy metals, pesticides, parental stress reactions and certain prescription drugs (Dietert, Dietert, & Dewitt, 2011). Although no one factor can be causally identified, definite differences in the specific methylation processes of PFC neurons postmortem between autistic individuals and controls have been identified (Cheung et al., 2010). Interestingly, although these biochemical differences were not readily apparent during early infancy, more precise inspection revealed subtle signs of a histone-lycine methylation marker (H3K4) beginning to spread in activation beyond areas found in non-autistic controls with further peaks in methylation identified in later PFC development. This mirrors the developmental progression of autism, with noticeable changes in behavior corresponding to changes in methylation patterns.

Executive Dysfunction and Disorders of Childhood

There is little disagreement that the self-regulatory behaviors associated with EFs are critical to academic and social achievement. As instruction moves from concrete to abstract and problem solving begins to involve multi-step procedures, the ability to think critically, attend to instruction, plan and organize an approach to novel problems, and retain and manipulate information become increasingly important. While overall cognitive ability remains the primary predictor of academic achievement, assessed skills such as inhibition, cognitive flexibility, and self-monitoring have been found to be almost as predictive of academic achievement in all core academic subjects (Latzman et al., 2010). Successful social interaction involves response inhibition and regulation, as well as the ability to maintain or modify behavior based on feedback (Lezak, 1982). The established role of the OFC in emotional regulation as well as previously mentioned well known studies on changes in behavior due to brain injury leave little doubt that executive dysfunction can have profound social effects (Damasio, 1994). Executive dysfunction occurs for a number of reasons and has been identified in various disorders of childhood. Pertinent research on the implications of executive dysfunction in clinical child populations highlights the importance of effective assessment.

ADHD. ADHD may have differing etiology depending on which of the subtypes is diagnosed. Self-regulation in children diagnosed with ADHD has been conclusively tied to hypo-activation of the dlPFC and anterior cingulate regions of the frontal cortex (Dickstein et al., 2006). Although the dlPFC circuit is most implicated in regulating attention, evidence exists that dysfunction in the OFC circuit is involved in hyperactivity/impulsivity (Hale & Fiorello, 2004). ADHD is considered a disorder of attention and inhibition, but there is also evidence that impairments in working memory cause an inability to use previous learning to alter behavioral patterns when faced with similar stimuli (Kofler et al., 2011). It is important to note that ADHD predominately inattentive type may not be related to executive dysfunction at all. Consistent with the "top-down" and "bottom-up" nature of cortical systems, daydreaming, confusion, passivity, and a lack of awareness can result from parietal dysfunction affecting attention through disorganization of sensory information affecting perception and interpretation of stimuli (Hale & Fiorello, 2004). This may explain why some children do not respond to ADHD medications that increase dopamine levels. Their symptoms, such as daydreaming, passivity, and failure to attend to relevant stimuli may be more a result of occipital – parietal impairments that would not respond to psychostimulant medication. Due to epigenetic factors, deficits in executive attention and control may also explain the higher prevalence of ADHD among low socioeconomic groups (Ardila et al., 2000). Deficits in executive attention control in impoverished, at-risk students may explain negative academic experiences that see levels of motivation decline over time (Howse et al., 2003). Although children from poverty exhibit high levels of motivation upon entering elementary school, they score more poorly on measures of sustained attention due to factors explained earlier. These difficulties result in more correction and greater difficulty with academic tasks.

Autism. Executive dysfunction is also linked to autism where deficits in working memory, planning, and switching rule sets occur in less structured tasks that require formation of strategies for successful task completion (Hill, 2004). Children with high functioning autism (HFA) or autism without significant cognitive impairments, show deficits in executive processes associated with concept formation and task planning in assessment and in flexibility, inhibition, and switching in observer ratings (Semrud-Clikeman, Walkowiak, Wilkinson, & Butcher, 2010). However, the etiology of these deficits may involve numerous cortical and subcortical circuits. Research has revealed neurological differences in children with autism such as smaller volume in the corpus collosum or excessive white matter in parietal and temporal areas. This may indicate that coordination of sensory and perceptual information may be inefficient in some cases, or, overwhelming in other cases. Due to an impaired ability to understand pragmatic and inferential language, verbal tasks requiring semantic organization of verbal information may also prove difficult, especially when such tasks have an executive component such as switching between responses. Common to all students with HFA are impairments in fluid reasoning and decision-making. These deficits have been associated with impairment in mirror neuron systems which allow for a mental representation of one's own actions or of the actions of others (Hadjikhani, 2007). This mental representation allows for analysis of another's goals and intentions. Emotional salience ascribed to these impressions is consistent with Damasio's SMH (Damasio, 1994). Impairment in the mirror neuron system may prohibit developing the emotional salience of certain responses to others actions, therefore impairing the development of social/behavioral routines that aide in planning and decision-making.

TBI. Almost all TBIs involve impairments in self-regulatory processes and severe TBI can impair regulation and goal setting abilities a decade after injury (Beauchamp et al., 2011). As mentioned previously, most tests of EF were first developed in studies involving brain injury. The results of TBI are difficult to predict especially in the developing brain. The brain undergoes many distortions after injury and the delicate nature of axonal connections is highly susceptible to acceleration and deceleration of the brain due to the impact effects of a coup-contracoup injury (Ewing-Cobbs, Barnes, & Fletcher, 2003). Hypoactivation of brain regions to facilitate healing can also have long-term effects on EF. Thus, TBI causes deficits in executive processes through diffuse cortical and subcortical dysfunction. Injuries that are specific to the central lobes; however, generally cause profound deficits in attention, goal orientation, inhibition, decision-making and perseveration as established by clinical samples in assessment (Delis et al., 2001a; Korkman et al., 2007b). As mentioned previously, developmental aspects of executive processes mean that some effects of TBI in children may not be evident until adolescence.

Learning disorders. Deficits in executive processes such as working memory have also been implicated the development of specific learning disabilities in math and reading (García-Madruga et al., 2014; Toll, et al., 2011). The more complex the math processes, the more important aspects of executive control, such as concept formation, cognitive flexibility and working memory, become. Deficits in reading fluency that involve functions of the PFC generally involve orthographic and not phonological difficulties (Hale & Fiorello, 2004). Children with reading difficulties may have attentional impairments, or problems with set maintenance, flexibility and sequential organization of responding. These issues manifest predominately as problems with fluency and retrieval speed in reading. Reading comprehension that is not impaired due to poor basic reading skills depends on an ability to simultaneously process the holistic "gist" of a passage (integration of information), while resolving discordant information such as idiomatic phrases, words with multiple meanings, or any subtle aspects of the text that do not integrate easily into the story narrative being processed. These are predominately functions of the left and right dlPFC respectively. A large part of building narrative is effective retrieval and comparison of acquired semantic knowledge, thus making working memory skills important for thorough analysis of complex literary works.

Summary

Connectionist theory helps explain normal brain development in terms of undeveloped neural systems that respond and mature in response to environmental demands. It also helps to define the processes by which executive functioning matures from an undifferentiated, predominately inhibitory system to a combination of differentiated but related functions. While connectionist theory provides a framework for the consistency of development, epigenetics illustrates how differences arise in the efficiency of executive processes due to environmental factors. Environmental risk factors such as poor living conditions as a result of low SES, as well as parent-child relationships, can literally change potential through gene suppression while mitigating factors such as early intervention and maternal care can enhance outcomes (Illig, 1998). With regard to assessment of childhood disorders, the interplay of multiple cognitive abilities and executive processes becomes blurred when analyzing disorders of cognitive and behavioral dysfunction. It is often impossible to separate the two even in disorders such as ADHD that are clearly associated with deficits in executive processes as dysfunction in subcortical units can mimic behaviors associated with executive dysfunction. Risk factors in development and differences in subcortical processes are some reasons why neuropsychological assessments seem to be less sensitive to differentiating disorders of executive processes than rating scales. Whereas assessment taps discrete skills, rating scales describe behaviors. As has been evidenced, behaviors associated with poor regulatory control can be the result of many factors outside of EF including learned behavior, difficulty with sensory integration, anxiety, and hyper-vigilance and deficits in development due to socio-economic risk factors. The confluence between such observer ratings and actual dysfunction in executive processes may lie in cases where the observer ratings of metacognitive dysfunction are clearly of clinical significance.

Purpose of the Study

The review of literature has sought to present the historical development of the theory of EF, outline the pertinent current theories of EF, and establish the importance of assessment. It has sought to describe and validate the types of assessment devices used to establish deficits in executive functioning and offer both theoretical and empirical validity for their use. Finally, it has sought to describe differences in adult and child models and how assessment results differ based on developmental age, risk factors, and clinical disorders. Overall, the review of literature sought to establish the importance of appropriate and valid means of assessing executive processes in children. Finally, this research study attempted to explore the validity of commonly used assessment instruments in supporting a factor structure in a child clinical sample as well as in predicting significant levels of observed executive dysfunction.

CHAPTER III

METHODOLOGY

The study was designed to discern if tasks associated with the measurement of EF would reflect valid latent constructs and then to use those constructs to predict observer ratings of cognitive executive dysfunction. The clinical sample and archival data presented the opportunity to explore whether assessment results would reflect parent ratings of significant difficulty with executive skills such as planning, organization, working memory, and initiation.

Participants

The data used in the analysis were archival data originating from neuropsychological assessments submitted as part of the KIDS, Inc. School Neuropsychology Post-Graduate Certification program. Assessment data and demographic information from comprehensive assessments conducted by student-clinicians over a period of 10 years were collated into a database from which the research sample was drawn. These evaluations were completed with school-aged children and adolescents who demonstrated significant academic, cognitive, and/or behavioral concerns. The vast majority of students were diagnosed with one or more of the following conditions: attention-deficit hyperactivity disorder (AD/HD), specific learning disabilities (LD), autism spectrum disorders (ASD), speech/language disorders, emotional disturbance, and medical disorders such as acquired brain injuries, tumors, and seizure disorders. The sample was limited to children between the ages of 8 years and 16 years, 11 months. This age range allowed for the inclusion of the common assessment tasks and mitigated the confounding developmental effects of early childhood.

Procedures

In addition to the documentation of informed consent, case studies submitted to KIDS, Inc. by candidates seeking certification in school neuropsychological assessment were required to contain notification that assessment data may be used for research purposes. Any evaluation containing documentation prohibiting use of the data for this purpose were excluded from the research database. To address the research questions set forth in the study, the database was sorted to facilitate identification of cases appropriate for the study. The primary sort for inclusion was the age of the participants, with the secondary sort noting the presence of a parent/guardian rating on the BRIEF. The remaining cases were then analyzed to ensure inclusion of at least 20% of the study variables and that participants with global deficits in cognitive ability were excluded. Finally, the data were scanned for outliers and assumptions of normality and homoscedasticity.

Because of the nature of archival data and the use of a wide variety of assessment instruments, missing data within a dataset may hamper analysis. Because of this, it was necessary to impute data within the individual datasets. Cases that contained at least 20% of the variables in the study and variables that contained at least 45% valid cases were included in the imputation model, which will be discussed in a later section. The imputed dataset was assessed to ensure it was representative of the original dataset. All further analyses were completed using the imputed data.

Measurement

The primary purpose of the study was to examine the factor structure and the predictive validity of direct EF assessment with a subset of children drawn from a large clinical sample. Test instruments chosen for the study contained tasks that have been empirically or theoretically validated as assessing one or more primary executive processes associated with inhibition/ initiation, switching/cognitive flexibility, working memory, or one or more facets of attentional control. All of the selected tasks were described in the review of literature including cancellation/matching, categorical reasoning, working memory, verbal and design fluency, and Stroop tasks. Variables were drawn from three assessment batteries: the D-KEFS, NEPSY-II, and the WJIII COG NU. Tasks were chosen to provide at least three variables in each task category. Final inclusion of variables depended on the availability within the sample, the likelihood of collinearality with other variables, and the percent of missing data within the cases being analyzed. In the second statistical analysis, the sample was divided according to parent ratings of EF, specifically the metacognitive index score from the BRIEF. The BRIEF is widely used in clinical practice and research to determine levels of cognitive and behavioral executive dysfunction (Gioia et al., 2002; Mahone et al., 2002). Table 1 provides a summary of the assessment tasks proposed for use in the EFA and reasons for inclusion/exclusion. Table 2 provides BRIEF scales and criteria for group membership.

Table 1

Variables Considered for Inclusion in the Exploratory Factor Analysis with Efficacy for Executive Function Assessment

Test –Subtest	Score Criteria and Type (Standard/Scaled)	Theorized Executive Constructs	Included/Excluded
D-KEFS			
Trail Making Condition 4	Completion Time	Inhibition, Attentional Switching	I
Color Word Interference Condition 3	Completion Time	Inhibition	E[1]
Color Word Interference Condition 4	Completion Time	Inhibition, Attentional Switching	I
Card Sorting	# Correct Sorts	Concept Formation, Cognitive Flexibility	I
Verbal Fluency Condition 3	Total Items Correct	Attentional Switching, Cognitive Flexibility	E[2]
Tower	Correct Designs/ #Moves	Monitoring, Planning, Working Memory	I
Design Fluency Condition 3	# Correct Designs	Monitoring, Attentional Switching	E[3]
D-KEFS 20 Questions	Total Correct	Planning, Cognitive Flexibility	I
NEPSY-II			
Animal Sorting	Correct Sorts	Concept Formation, Cognitive Flexibility	I
Inhibition: Inhibition	Combined Time/Errors	Inhibition	I
Inhibition: Switching	Combined Time/Errors	Inhibition, Attentional Switching	E[4]
Word List Interference Recall	Number Correct	Working Memory/Register Updating	E[4]
Auditory Attention	Combined Correct/Errors	Sustained/.Selective Attention	E[3]
Response Combined	Combined Correct/Errors	Selective Attention, Attentional Switching	E[3]
WJ III COG NU			
Concept Formation	Total Correct	Cognitive Flexibility, Concept Formation	I
Analysis Synthesis	Total Correct	Focused Attention, Working Memory	I
Numbers Reversed	Total Correct	Working Memory	I
Auditory Working Memory	Total Correct	Working Memory, Selective Attention	I
Visual Matching	Total Correct	Selective/Sustained Attention	I
Decision Speed	Total Correct	Selective/Sustained Attention	I
Pair Cancellation	Total Correct	Selective/Sustained Attention, Monitoring	I

[1]Excluded due to sampling inadequacy after initial EFA. [2]Excluded due to Heywood Case after initial EFA [3]Excluded due to lack of correlation with dataset, prior to initial EFA [4]Excluded prior to data imputation due to low occurrence in sample

Table 2

BRIEF Scoring Criteria for Group Membership in the Logistic Regression

Component Scales	Criteria	Group Placement
Initiate, Working Memory, (ED) Monitoring, Planning,	Metacognitive Index T-Score $\geq$70	Executive Dysfunction
Organization of Materials (NOED)	Metacognitive Index T-Score $\leq$ 60	Executive Competency

Explanation of Reliability and Validity

Reliability is the degree to which an instrument can produce the same results under the same conditions over time without substantial variance (Field, 2009). Reliability or the precision of a given measurement can be established in various ways. The most common is to establish the test – retest reliability of the measure by assessing the same group of people twice within a given time frame under the same conditions. The linear relationship between the test and retest is expressed as the Pearson Product Moment correlation, denoted by “r”. Due to practice effects, alternate forms of a test are sometimes used for re-administration. Internal consistency represents another measure of reliability that ensures the items within the scale are related and consistent with one another.

One method of assessing internal consistency is the split half method, which involves parsing the subtest items into equal halves. The resulting correlation is expressed as Chronbach’s alpha statistic (r11) measuring the covariance between item pairs on the test. A third type of reliability measure, decision-making consistency

(DMC), was utilized in the NEPSY-II battery. This method used data from the first administration of a task to predict the resulting category into which a second administration would fall. DMC measures the broad consistency of the instrument in correctly classifying an outcome. The proportion of correctly classified scores on the second administration corrected for measurement error, or Chronbach's kappa, is the measurement of reliability. The method is based on the premise that individual variations in test items are only problematic if they significantly changed the expected overall results the instrument provides. Test reliability is a key component of measuring the criterion and content validity of the instrument; that is, does it consistently measure the construct it purports to measure and do the items represent the construct being measured (Field, 2009). Construct validity, or whether the construct itself has been validated by previous research, is necessary for the test or instrument to have meaning. The more established the construct, the greater opportunity to prove both convergent or discriminant validity or the ability for an instrument to correlate with or discriminate between other instruments that measure facets of the specific construct. Predictive validity indicates the degree to which a measure can predict outcomes consistent with those expected or the extent it can identify suspected differences in the given construct between groups. For instance, an instrument shows predictive validity in measuring attention if it is proven to differentiate between a control group and a clinical group. Moderate to high correlations with measures that have been established to measure the same construct provide convergent validity for the measure while lower correlations reflect the ability of the instrument to discriminate between constructs.

Delis-Kaplan Executive Function System (D-KEFS)

The D-KEFS is a comprehensive measure of executive functioning that contains numerous traditional neuropsychological tasks, such as the Trail Making task, the Stroop test, and the Tower of Hanoi (Delis et al., 2001b). The D-KEFS combines various tasks into one instrument standardized on children, beginning at 8 years of age through to adulthood. The D-KEFS does not have a strong theoretical base for the inclusion of its tests and, consequently, does not purport to have a factor structure. The inclusion of tasks is based upon the predictive validity of the tasks to identify frontal lobe injury. True to the training of its authors, the D-KEFS uses a process approach that seeks to standardize qualitative process data (e.g. number of corrected or uncorrected errors, number of task rule violations). The D-KEFS was designed to have multiple trials of increasing complexity to assess the efficiency of lower level skills such as processing speed. For instance, the Trail Making task includes five trials. The first trial tests simple visual discrimination by requiring identification of a target among distracter targets. The second and third trials require simple sequencing and visual scanning skills connecting dots containing only letters or numbers. The fourth trial assesses inhibition and switching while the fifth trial involves simple paper and pencil tracing speed. By comparing results on all of these items, a more in-depth assessment of the role of executive processes in any deficits is possible (Delis et al., 2001b).

Specific tasks from the D-KEFS. The following tasks from the D-KEFS were chosen based on the following criteria: empirical validity, construct validity, commonality of use in the available database, and the reported skills being assessed.

Trail Making - Condition 4. This fourth trial of the D-KEFS Trail Making test requires the participant to connect dots alternating between letters and numbers in sequential order. This paper and pencil task is printed on a 17 X 22 inch sheet. When the examinee makes an error, the examiner halts the task, marks the error, and then allows the examinee to continue; thus, there is a time cost for each error made. The task yields scaled scores for both total time to completion and number of errors.

Card Sort. This task requires the examinee to sort eight cards into groups of four based upon a common element in each group. Each card is printed with shapes or lines and a word, and can be sorted either by a visual-spatial trait (color) or by a verbal trait (word category). The examiner presents an example sort and then directs the examinee to create a novel sort and explain the common trait in each of the two groups. After a successful sort, the examinee is asked to develop another grouping. The task ends when a time limit is reached or the examinee decides he or she can no longer generate any sorts. Although there are many supplementary scores available, such as the number of novel sort errors, incorrect descriptions of sorts, and repeated sorts, the two primary measures are correct sorts and correct verbal descriptions of sorts. Scaled scores represent both primary measures.

Verbal Fluency – Condition 3. The initial two conditions in the D-KEFs verbal fluency subtest ask the examinee to name as many words as possible given a specific cue within a one minute time period. There are general rules presented in each of the two trials. Condition 1 involves recall of words that begin with a certain letter. Condition 2 involves recall of words based on a category. Condition 3 requires the examinee name items from two different categories while alternating between the categories. The scaled score used in this research included the total number of correct words given by the examinee.

Color-Word Interference Conditions 3 and 4. This task represents an update of the Stroop task. Condition 1 requires the individual to correctly state the color of small colored bars arranged in rows on a page under timed conditions, while Condition 2 requires the individual to read the printed names of colors arranged in rows on a page, also under timed conditions. Condition 3 assesses inhibition by having the color names printed in a color than does not represent the word (e.g. the word "red" printed in green ink). The participant is asked to name the colored ink and not read the words. Condition 4 involves switching between reading the word and naming the color ink based on a given cue. The total time to completion on the third trial yields a scaled score. The total time and total errors on the fourth condition also yield scaled scores.

Tower. Based on the Tower of Hanoi, this task involves moving two to five rings of differing size between three pegs to replicate a pictured example. The participant must follow certain rules: a larger ring cannot be placed on a peg that contains a smaller ring, and only one ring can be removed from a peg at a time. To be successful in the completion of more difficult items, an examinee must use planning and working memory skills to detect the pattern necessary to accomplish the task. The primary score for this task, represented as a scaled score, is the total number of correct sorts scored with extra points given for completing the task in as few moves as possible within the time limit.

Design Fluency- Condition 3. This task requires that the examinee construct figures by connecting small arrays of dots using four lines under timed conditions without repeating any designs. In Condition 1, the examinee connects black dots. In Condition 2, there are also white dots, which the examinee must ignore. In Condition 3, the examinee must switch between black and white dots. The primary score for this task in the number of correct of non-repeated designs.

NEPSY-II, A Developmental Neuropsychological Assessment, Second Edition

The NEPSY-II is a compilation of individual tasks designed to assess neuropsychological function in children ages 3 to 16 years (Korkman et al., 2007a). The 36 tasks assess 6 content domains: Attention and Executive Functioning, Language, Memory and Learning, Social Perception, Sensorimotor, and Visuospatial Processing. Although these domains do not constitute statistically validated factors, each subtest was chosen for its ability to tap a variation of the primary skill domain in which it was included.

Similar to the D-KEFS, the NEPSY-II manual reports low to moderate test-retest reliability for most tasks involving executive functioning. Additionally, the NEPSY-II also provides reliability coefficients for decision-making consistency. This method uses response patterns on the first administration of a task to predict the likely range of results on a second task. Using this measurement, the NEPSY-II tasks show moderate to high test-retest reliability. Similar to the D-KEFS, the NEPSY-II tasks of Attention and Executive Functioning show low to moderate correlations with one another due to the different aspects of executive skills measured.

The NEPSY-II also reported results derived from small clinical samples administered between 22 and 32 of the subtests included on the final battery. These groups included Traumatic Brain Injury, Asperger's Disorder, Intellectual Disability, Language Disorder, Reading Disorder, ADHD, Deaf and Hard of Hearing, and Emotionally Disturbed. The studies compared these clinical groups to a control sample matched on age, parent education level, race/ethnicity, and sex. The results indicated that the NEPSY-II domains of Attention and EF and Language were the most sensitive to predicting differences between typically developing children and the clinical samples.

Specific NEPSY-II tasks in the research. The domain of Attention and Executive Functioning is comprised of tasks that assess several constructs that share common executive processes measured in distinct ways (Korkman et al., 2007b). Six tasks measure selective and focused attention, planning, self-monitoring, inhibition of automated responses, cognitive flexibility and categorical thinking, and the ability to switch between rule sets in a task.

Animal Card Sorting. This is a categorical card-sorting task consisting of eight cards depicting animals in various terrains and environmental conditions. The examinee is asked to sort the cards into two distinct groups of four cards with each set having a common trait differentiating it from the other set. Once a sort has been presented, the examinee is asked to continue to develop novel ways of sorting the cards until a time limit is reached or the examinee indicates that he or she cannot generate further sorts. Unlike the D-KEFS, the NEPSY-II card sort does not score the examinee's ability to describe the basis for the sort. Like other tasks measuring cognitive flexibility and concept formation, this subtest encompasses a wide range of executive processes including task initiation and self- monitoring (Brooks, Sherman, & Strauss, 2009). The task yields a scaled score indicating the amount of correct sorts.

Auditory Attention and Response Set. Meant to assess facets of attention, inhibition and set shifting, this subtest consists of two separate tasks (Korkman et al., 2007b). Auditory Attention assesses both sustained and selective attention by requiring the child to touch a specifically colored circle every time he or she hears that color spoken among a list of words. The child must also ignore other colors named, and inhibit touching other colored circles. Response Set comprises the second half of this task. In this activity, the examinee is instructed to touch a colored circle other than the color being named (e.g. touch yellow circle when you hear red and vice versa), as well as touching the same colored circle as stated. This task is designed to assess the ability to switch between rule sets and inhibit an automated response while sustaining attention. The Auditory Attention Response set trials yield scaled scores combining the number of correctly identified targets and the number of commission (missed target) errors.

Inhibition-Inhibition and Switching. The Inhibition subtest is another Stroop task that uses shapes and arrows in lieu of words and colors (Korkman et al., 2007b). It includes two separate tasks each involving three trials. In the first task, the child must rapidly name rows of shapes, then state the name of the opposite shape, and, finally, alternate between either the shape or its opposite based on whether the shape is black or white. In the second task, the child must first name the direction of rows of arrows, then the opposite direction the arrow is pointing, and finally the correct or opposite direction based upon whether the arrow is black or white. The first trial of each task measures rapid naming of the objects, the second trial measures inhibition of habitual responses, and the third trial measures cognitive switching. The final score is the combination of the completion times and number of errors on the tasks.

Word List Interference. Word List Interference assesses verbal working memory in a unique way. The examinee is not required to manipulate data but, instead, is asked to listen to a set of two word lists of varying length. After each set is presented the examiner repeats back, in order, the words he or she has heard. The examinee is then asked to recall the words on the first list presented and then the second. This constitutes the recall portion of the trial, which is the most direct assessment of the facet of working memory that requires one to discard or set aside one set of information to accurately recall another. Word List Interference belongs to the Memory and Learning domain; however, as mentioned previously, working memory is theorized to involve executive processes in different proposed models (Miller, 2013; Miyake et al., 2012), which is why the recall score from this subtest was selected for inclusion in this study.

Woodcock Johnson III Tests of Cognitive Ability, Third Edition, Normative Update (WJIII COG NU)

The Woodcock Johnson cognitive test battery was the first widely published assessment battery based upon CHC intellectual theory and its standard and extended batteries cover a wider range of CHC broad and narrow cognitive abilities than any other major assessment device (McGrew et al., 2007). These abilities are weighted to derive an overall intelligence quotient called the Global Intellectual Ability (GIA). The cognitive abilities measured by the WJIII COG NU include fluid reasoning (*Gf*), comprehension-knowledge (*Gc*), long-term storage and retrieval (*Glr*), processing speed (*Gs*), visual-spatial reasoning (*Gv*), auditory processing (*Ga*), and short term memory (*Gsm*). As mentioned in the review of literature, *Gf* or fluid reasoning can be equated with some theorized executive processes such as conceptual thinking and cognitive flexibility. In addition, there are two subtests measuring working memory.

The WJIII COG NU was normed using a sample of 8,782 subjects representative of the 2000 United States census (McGrew et al., 2007), and the normative update was a realignment of the data in the original sample to better match the 2005 census. All tests comprising the WJIII COG NU battery were normed on a diverse set of participants from groups representing the United States population by geographic location, age, race/ethnicity, and SES. The WJIII COG NU can be considered to have strong psychometric properties because it was designed based upon extensive research within an established theory of intelligence. The scoring uses sample data to weight individual items on each subtest based on the likelihood that individuals at a given age and grade level will answer the question correctly. This allows for both normative referenced scores and criterion-referenced scores (McGrew et al., 2007; Schrank, Miller, Wendling, & Woodcock, 2010). The theoretical model provides strong evidence for the content validity of the test. Criterion validity was established by extensive comparisons of the GIA with six other instruments measuring general intelligence ($r > .70$; Schrank et al., 2010).

The WJIII COG NU was also administered to clinical samples to identify sensitivity to specific disorders of childhood such as Autism, TBI, Reading disorders, Mathematics disorders and AD/HD (Woodcock et al., 2007). For most groups, standard scores generally fell below the sample mean of 100; however, with the exception of processing speed scores for the Autism and TBI samples, standard scores remained within the average range (SS > 90). The Language disorder group was the only clinical sample who exhibited below average scores in multiple cognitive domains.

Tasks from the WJIII COG NU. The following tasks were considered relevant to the research because they were variations on the categories of assessments deemed to tap executive processes. All scores on the WJIII COG NU are reported as standard scores with a mean of 100 and standard deviation of 15.

Concept Formation. This task is a variation on category tasks and assesses inductive reasoning skills. The participant is asked to determine why one or more shapes of various sizes, pairings, and colors have been grouped together while others have been excluded. It requires rule application and frequent switching of one rule to another. Corrective feedback is provided on all but the last few items but is limited to providing the correct answer.

***Analysis-Synthesis*.** In contrast to concept formation's assessment of inductive reasoning skills, Analysis-Synthesis assesses sequential, deductive reasoning skills. A code key is presented that indicates what combinations of colored squares equate to a single colored square (e.g. black square and blue square = one yellow square). The participant is then asked to solve equations by filling in blank squares with the correct color square. As problems become more complex, two steps of calculation are often necessary to solve problems.

***Visual Matching*.** The processing speed domain of the WJIII COG NU assesses the ability to accurately and efficiently make comparisons based on rapid visual searches. Visual Matching requires rapid selection of identical numerals in rows of numerals under timed conditions. The task measures not only perceptual speed but selective and focused attention. As numbers get larger (from one digit to three digits in length), the task also requires identification of an effective strategy to proceed with efficiency.

***Pair Cancellation*.** This processing speed task requires quick identification of a set of two pictures in specific sequence. The targets appear on a page in rows amongst non-target pictures. The task assesses vigilance or sustained attention, as well as selective attention. An effective strategy is also necessary to identify targets efficiently.

Decision Speed. Decision speed creates a slight variation on simple matching tasks by requiring symbolic comparison that assesses speed of semantic reasoning. The participant is asked to select two pictures in rows of pictures that are most alike in some way. The pictures are not identical but have some categorical trait in common that separates them from the remainder of the pictures. The addition of a reasoning component to this task creates an additional cognitive load beyond simple matching.

Numbers Reversed. The domain of short-term memory is assessed on the WJIII COG NU by three different tasks. Of these, two require the manipulation of information instead of just rote memorization and repetition. Numbers Reversed is one of these tasks. It requires the participant to listen to a series of numbers, and repeat them in reverse order. The reordering of information in short term memory taps auditory working memory.

Auditory Working Memory. This task is the second in the memory cluster to tap working memory skills. It requires the participant to listen to a combination of words and numbers, then repeat them stating the numbers first from low to high and the words in alphabetical order. The raw score points are awarded for correctly ordering the numbers and words.

Behavior Rating Inventory of Executive Functioning (BRIEF)

The BRIEF (Gioia, et al., 2000b) is an observer rating of EF, from which two composite indices, metacognition and behavioral regulation, are derived. These indices factor into an overall global scale called the Global Executive Functioning Scale. Each composite scale is comprised of subscales. The Initiate, Working Memory, Organization of Materials, Planning/ Organization, and Monitoring scales compose the Metacognition Index, while the Inhibit, Shift, and Emotional Control scales combine for the Behavioral Regulation Index. The individual scales and composite scales are T-scores, in which scores above 70 are evidence of significant impairment. The normative data and factor structure of the BRIEF were obtained from samples of 1419 parents (815 girls and 604 boys) and 720 teachers (403 girls and 317 boys) stratified based on United States census data for ethnicity, socioeconomic status and population density (rural, suburban, and urban). The validity of the BRIEF has been questioned due to the majority of the normative sample being drawn from a single geographic area: the state of Maryland. The authors have since conducted a meta-analysis of 39 published research studies from various regions of the United States and Canada that included BRIEF data on 2539 typically developing children (Roth, Erdodi, McCulloch, & Isquith, 2015). From this data, they concluded differences based on geographic area are minute (≤ 1 T-score unit difference) in all indices and in the global score. Mean internal consistency ratings for the BRIEF reported for clinical populations mirrors the normative sample and range from .82 to .98. Only moderate correlations have been found between teacher and parent ratings.

The BRIEF was used in this study to divide participants into two distinct groups of average and impaired executive functioning. Group membership was determined by parent ratings on the Metacognitive Index of the BRIEF. The Metacognitive Index score was chosen due to the emphasis on metacognitive problem solving skills rather than general behavioral regulation. In addition, the factor structure of the Metacognitive Index was better preserved than the behavioral regulation index in clinical samples. To avoid error due to confidence intervals and standard error of measurement, students with T-scores greater than 70 were included in the impaired group. Students with T-scores less than 60 were identified as the non-impaired group.

Data Analysis Procedures

Descriptive Statistics

Statistics were analyzed by demographic variables such as age, gender, and ethnicity to ensure there were limited differences among groups for the study variables. For instance, a statistically significant difference in test scores on more than a few variables based on gender would require statistically weighting scores to ensure equal representation in the analysis. To ensure the validity of statistical analyses, certain preliminary issues were addressed. If datasets contained extreme scores, were not normally distributed, had missing data points, or were represented by different types of data (e.g. categorical, ordinal, continuous), problems could arise that threatened the validity of statistical analyses.

Outliers

Outliers are extreme or even impossible data points in a variable that confound results. To identify impossible or highly unlikely values in the original dataset, box and whisker plots were used to provide a quick visual indication of data points that were significantly beyond the range of other data points. After imputation of data took place, the dataset was again scanned for significant outliers. Because most variables had substantial numbers of scores falling at extreme ends of possible scores, it was deemed inappropriate to exclude a large amount of variables that represented valid scores within the dataset. Instead, it was more important the scores across variables were consistent within each case. A statistical method of eliminating inconsistent data points involved the calculation of Mahalanobis distances from the mean for each variable by case (Mertler & Vannatta, 2005). This statistic was obtained through linear regression of the variables with case numbers as the dependent variables. By comparing this result with a Chi-Square statistic with degrees of freedom (df) equal to the number of variables in the study, inconsistent results within each case were identified.

Normality

The assumption of normality of data sets states that it is assumed that data points in each variable follow a normal distribution. Datasets that are highly skewed can affect the validity of statistical analyses especially with smaller sample sizes. Each variable for the current study was analyzed using a statistical analysis called the Kolmogorov-Smirnov test, which compares the distribution of variables against a standard normal distribution. If the result was statistically significant, skewness and kurtosis of each variable distribution were calculated to ensure tolerance limits for these analyses were not exceeded. Skewness levels between +/- 1 and kurtosis levels between +/- 3 are identified as indications of normality which can be assumed without transformation of the variable.

Multicollinearality

Bivariate correlations were calculated to determine the correlation of each variable compared to every other variable (Field, 2009). This correlation matrix was analyzed to ensure multicollinearity, or variables that correlate highly, did not threaten the assumptions of the statistical analyses. In addition, the matrix served to reveal which variables were substantially unrelated and unlikely to be appropriate for factor analysis. Generally, if a variable failed to show moderate correlations with other variables ($r \geq .30$), the variable was removed from analysis (Yong & Pearce, 2013).

Missing Data

Missing data were expected due to variations in tests used across the sample. The missing data were addressed by imputing missing data points. Imputation of missing data does not “predict” an individual’s score on a given variable, rather it offers better estimates of the population parameters based upon various factors from both the sample and individual (Weiner, Schinka, & Velicer, 2012). In general, data can be missing completely at random (MCAR), missing at random (MAR), or missing systematically. Variables MCAR have no pattern or explanation for missing values and are unrelated to the target variables being assessed. This is a very unlikely situation. MAR patterns are the most likely explanation for any missing data not related to a systematic characteristic of the study. The assumption for MAR is met if the pattern for missing data is not related to a specific variable responsible for the missing data points. Due to the differences in the frequency of use of different test instruments among the examiners, the MAR assumption is the most likely explanation for the patterns of missing data found.

In data imputation, it is important to consider the amount of missing data. As mentioned, only cases with at least 20% of the study variables and variables with less than 55% missing data were included in the analysis. Once variables were chosen, a method of imputation was established. Expectation Maximization (EM) has been established as a reliable imputation method. This is an iterative process that uses the sample statistic for the variable in question to predict a value for the missing data (Osborne, 2013). However, single imputation using EM is recommended only in cases of small amounts of missing data. With large amounts of missing data, multiple imputation (MI) using the Monte Carlo Markov Chain (MCMC) procedure is recommended to enhance the integrity of imputation results (Hurley & Scandura, 1997). This process uses boot strapping (replication of random datasets based on sample parameters) to impute multiple possible samples of the theoretical complete dataset, while still representing a random value that is valid within the parameters of the dataset. To ensure accuracy, 10 imputations of the dataset were drawn using the MCMC procedure. The means of these random imputations were then averaged to create the most likely value for the missing variable. The new dataset was then analyzed to ensure it was representative of a theoretical original dataset in which no missing values occurred.

First Statistical Analysis – Exploratory Factor Analysis

1. *Is there a discernible factor structure of EF in a clinical sample of children that emerges from commonly used assessment instruments in detecting executive dysfunction? If so, how many factors best describe the data and do these factors represent similar constructs reported in research on adult and child populations?*

The appropriate analysis to answer the first research question was a factor analysis. This statistical method determines shared variance among multiple predictor variables and then groups them into hypothetical or latent variables called factors (Mertler & Vannatta, 2005). The purpose of factor analysis is to both reduce the number of variables in an analysis and reveal underlying constructs measured by the variables (Brown, Hendrix & Hedges, 2011). Factor analysis reduces statistical error in further analysis by substantially reducing the number of variables in the analysis and eliminating variables unrelated to others variables or variables whose high correlations make them redundant.

Different methods of factor analysis exist; however, given the presenting data, an exploratory factor analysis (EFA) was deemed the most appropriate. When a researcher is certain that data will conform to a specific model, a confirmatory factor analysis (CFA) is the better choice. While variables are allowed to load freely on different factors within an EFA, variables in a CFA are constrained to load on specific constructs in the a priori model (Cramer, 2003). Although CFA is a powerful modeling technique for extraction of factors, it requires the researcher to start with a strong expectation that the variables chosen will load on specific theorized factors. When the researcher is less sure, it is not wise to place model restraints upon the variables; rather, an EFA is a more appropriate statistical technique.

The most common form of data reduction is principal components analysis (PCA), which is often used in analysis of latent variables (Costello & Osborne, 2005). PCA simply reduces a large amount of variables to a smaller number of components that explain all of the variance from each variable occurring in the original variable set (Field, 2009). While a PCA accounts for more variance, it generally produces inflated results, as it assumes no error in the measurement of the original variables. The PCA, therefore, makes no assumptions regarding common underlying factors. True EFA requires that error is assumed to be inherent in the study variables and seeks to estimate shared versus unique variance in the variables prior to how they group together (Chen, 2010).

By nature, EFA is an intuitive process that requires multiple levels of analysis as to how variables factor together. EFA does not provide a definitive answer to the relationship between variables; rather, it suggests underlying factors derived from different combinations of the data (Costello & Osborne, 2005). EFA analysis is predicated on justifiable decisions based on both statistical indicators of the validity of emergent factors and the researcher's knowledge of prior research. The mathematical method of factorization used can depend on the distribution, consistency and quality of the data as well as what questions are being answered by the research.

There are different methods of EFA based on ways to derive eigenvectors or starting values for regression analysis of how the dependent variables relate to the independent underlying, non-observable factors (Taherdoost, Sahibuddin, & Jalaliyoon, 2004). Research has generally shown that, although there are advantages in analysis, the results produced by different methods should be extremely similar. In fact, significant differences in results from one method to another may indicate a flaw in the researcher's conceptualization of the data. Common methods used in EFA include Principal Axis Factoring (PAF), Maximum Likelihood Analysis (MLA), and Unweighted Least Squares (ULS; Chen, 2010). Principal Axis Factoring (PAF) is a method of factor analysis that is robust to violations of the normality assumption and is commonly used when results are not generalized to a larger population. It is designed to mathematically capture as much variance as possible in the model in the first extracted factor, thus maximizing the relationship between variables with the strongest relationships.

Although PAF is considered the default analysis for EFA, Maximum Likelihood Analysis (MLA) is superior in many ways. This method has the advantage of allowing for various measures of determining how well the results in a sample fit a theorized population model. It is generally used in CFA analysis to assess an a priori theory; however, it can also be applied to EFA, specifically when attempting to determine the number of valid factors on which the variables are allowed to load (Preacher, Zhang, Kim, & Mels, 2013).

The MLA procedure is heavily dependent on the assumption of normality to compute the reproduced correlation matrix. Small sample sizes and other factors that cause fluctuation in the dataset can result in negative values in the reproduced matrix which is a primary cause of Heywood cases (Kolenikov & Bollen, 2012). Heywood cases can also be caused by groups of highly correlated variables in the reproduced correlation matrix of the EFA. The way MLA calculates eigenvectors seems especially susceptible to generating variables with negative error variance which may allow a variable to load on a factor with a communality that exceeds 1.0. When the analysis attempts to compute the square root of a variable with negative variance the extraction is terminated. When an MLA analysis is mathematically untenable, comparable methods of analysis such as Unweighted Least Squares (ULS) can be used. ULS attempts to minimize unexplained variance in the model and has been shown to perform better than MLA in estimating population parameters when communalities among variables are not as strong (Coughlin, 2013). This fact allows for the use of the same goodness of fit indices reflecting model acceptability as in MLA. In the current analysis, MLA was used as the default extraction method; however, the MLA failed to generate a solution because the correlations between variables and factors were negative. Thus, the final analysis was completed using ULS, which avoided the mathematical anomalies caused by the estimation of the covariance matrix.

Prior to the extraction of factors in the EFA, a correlation matrix was generated to identify if variables correlated with one another to create latent variables and if variables exhibited multicollinearity or, in effect, measured exactly the same thing. The correlation matrix was generated as part of the descriptive statistics. The correlation matrix produces a statistic called the determinant, which is a measure of the area of the plotted data and values above .00001 indicate multicollinearity should not be a concern (Cramer, 2003). Bartlett's Test of Sphericity was used to ensure the correlations between variables were not poor. The Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) was used to analyze the proportion of variance between variables that may represent common factors. In addition, the anti-aliasing matrix was generated to examine the measure of sampling adequacy (MSA) for each individual variable. KMO values above .70 were an indicator that the variables adequately reflected the factors extracted. Individual variables with MSA <.50 were excluded from further analysis (Field, 2009).

Once the validity of the data was established, the criteria for selection of the number of factors could be determined using Eigenvalues (Mertler & Vannatta, 2005).

Eigenvalues are a mathematical measure of the dimensions of correlated data on a scatterplot. The higher the Eigenvalue of a group of data, the more those variables are clustered together and on the same linear vector (Field, 2009). The number of derived Eigenvalues equals the number of variables; however, only those factors that explain a substantial amount of variance were included in the final analysis. The cut-off for Eigenvalues to be retained was specified a priori. Kaiser's rule states that Eigenvalues above 1 (e.g. those that capture 1% of the overall variance for each variable in the dataset) should be retained.

Although Kaiser's rule is widespread in use, it is considered to be the least valid method of factor retention, especially for samples below 300 (Costello & Osborne, 2005). Cattell's scree plot is considered a better method of deciding upon factor retention. The analysis of the scree plot graph can reveal when Eigenvalues begin to explain less and less of the variance thus creating a "leveling off" effect in the graph. A scree plot analysis is, however, not always definitive. If the true purpose of an EFA is to determine whether variables group together in a specific way that is greater than chance, then it stands to reason that any factor chosen should have an Eigenvalue that occurs at a rate greater than chance (Watkins, 2006). To do this, it is necessary to run multiple simulations of sets of random data based upon the number of variables and participants in the actual study. Random data are generated, and a correlation matrix is produced. A PCA is then run for the data listing the Eigenvalue for the number of factors corresponding to the number of variables in the study. This process was repeated for a number of iterations of random data and the mean Eigenvalues for each factor were reported. Eigenvalues for the actual data can be compared at each level to the mean Eigenvalue of the random datasets. If it is greater in value, the data results occurred at a rate greater than random chance; however, if it is less, than it is possible the grouping occurred as a result of random chance. This technique is known as PCA parallel analysis (PA), and, although the method has existed for some time (Watkins, 2005), modern computer software programs allow for the rapid generation and analysis of up to a thousand random data sets simultaneously. To perform multiple random runs of data to determine all likely outcomes is a process known as a Monte Carlo simulation. This process can be conducted with various software packages to obtain Eigenvalue "cut-off" scores. To determine appropriate factor extraction, a combination of scree plot analysis, and PA was used.

To maximize interpretation of the EFA, a factor rotation (Field, 2009) was utilized. The axes upon which the data is plotted are rotated to minimize the variables' ability to correlate with other factors. Two types of rotation can be employed. Orthogonal rotation prevents any correlation with more than one factor while oblique rotation allows factors to correlate. Oblique rotation is used when there are theoretical reasons to believe that the factors are related. In this research, it has been shown that depending on age, the individual variables may tap more than one executive construct; thus, oblique rotation was selected. Direct oblique rotation is the recommended form of oblique rotation for analysis of most datasets. Once factors are extracted and the percent of variance explained by them are established, a table of communalities is produced showing the proportion of each variable's variance that can be explained by the retained or extracted factors.

The factor loadings of each variable are given in the rotated component matrix (Field, 2009). Oblique rotation generates a pattern matrix, which shows the degree to which variables load onto specific factors, and a structure matrix, which takes into consideration the correlations between factors. The structure matrix allows for the interpretation of how interrelated the factors are. For instance, if multiple variables load substantially onto more than one factor, then the variables most likely measure multiple constructs equally.

Determining the best factor representation of the variables. "Goodness of fit" statistics are used to determine if the factor data fits estimated population parameters. Although goodness of fit statistics are usually reserved for CFA to decide between competing theoretical models, these measures can also indicate the degree to which the results of the EFA analysis fit the data (Hooper, Coughlan, & Mullen, 2008). The Comparative Fit Index (CFI) is a measure recommended for smaller samples that contrasts the resulting model with a baseline model assuming zero order correlations between variables. The higher the CFI, the more variance in correlation is explained by the data. Generally, a CFI >.90 is considered adequate. The Non-Normed Fit Index (NNFI) also known as the Tucker-Lewis Index, is similar to the CFI except it does not adjust for degrees of freedom meaning it does not skew results due to the parsimony of a model. Two other measures of fit concentrate on the amount of residual variance not explained by the data, rather than comparison to a theoretical model. These are relative fit indices as opposed to the NNFI and CFI, which are considered absolute fit indices. The Goodness of Fit (GFI) index simply compares the total variance in the reproduced covariance matrix to the total unexplained variance. Values above .90 indicate the model explains the relationship among variables adequately. Another measure of fit based on covariance residuals is the Root Mean Square of Residuals (RMSR). This number represents the absolute value of all of the co-variance residuals. According to Hattie (1985), a long standing measure of goodness of fit called Kelley's Criterion, has been whether the RMSR exceeds the standard error of residuals as a function of the sample size. This criterion for the sample used was calculated as .075. Values below this amount for any factor model indicate an acceptable fit to the data.

Analysis of necessary sample size. For factor analysis, a "rule of thumb" noted in research is that there should be 10 sets of data for every original variable in the analysis (Costello & Osborne, 2005; Mertler & Vannatta, 2005). More specifically, necessary sample size depends on the degree of communality. High communality assures a good fit to the theoretical population parameters and greatly reduces the impact of low sample size, overdetermination or identification of too many factors and model error (MacCallum, Widaman, Preacher, & Hong, 2001). In contrast, low communality may require much more than the prescribed 10 data sets per variable for a valid solution to emerge. Initially, 20 variables were proposed; however, variables were eliminated in the analysis for various reasons including excessive missing values, a lack of moderate correlations with other variables, and inadequate MSA.

The determinant factor for an appropriate sample is again, the communality of variables. If only factor loadings above .50 are considered valid, then a sample of 100 is adequate for analysis; however, if factor loadings above .40 are deemed acceptable, then a sample size of approximately 200 is required (Hair, Black, Babin, & Anderson, 2010). Based on this data, factor loadings of .40 or more were considered interpretable. In the case of complex loadings (loadings on more than one factor of .32 or more), deference was given to the first factor upon which the variable loaded. This is because the initial factor captures the most variance in the model. Once the latent variables were identified, the common thread of the variables that have acceptable factor loadings were examined to help identify the construct or constructs that are most likely being assessed.

Second Analysis - Logistic Regression

2. *Can the latent variables derived from factors in the first analysis predict observer reported dysfunction in cognitive executive processes such as working memory, organization, planning, and self-monitoring?*

One of the primary issues in research between observed and assessed EF is that the correlation between the two is not consistent over all levels of functioning. In other words, fluctuations in impairment or efficiency on assessment instruments that fall within normal parameters do not correlate to fluctuations in impairment or efficiency in observed behaviors. As mentioned previously, these phenomena may result from factors such as differences in efficiency in core cognitive processes and specific instruction in compensatory strategies to remediate EF deficits. A lack of sensitivity in assessment devices may also exist due to skewed normative data where wide ranges in performance result in small changes in standard scores or where reliability coefficients are not sufficiently high. What is evidenced in the research is that assessments of EF are highly sensitive to significant dysfunction and disruption of cognitive regulatory processes, especially in cases of TBI (Ardila, 1992; Spencer-Smith & Anderson, 2009). Thus, although direct correlations between tests and observations of EF have not been shown in literature, there should be a stronger relationship in the ability of assessments to predict observed executive dysfunction. If this holds true, the latent variables developed in the EFA should accurately predict group membership between normal and impaired EF, which is the impetus behind the second research question.

To analyze a set of continuous variables as a predictor of a dichotomous outcome variable, a logistic regression (LR) is the prescribed statistical analysis (Field, 2009). LR is used when a group of predictor variables are used to explain a dichotomous outcome. Simply put, it tests the likelihood that changes in the value of a given independent variable x will result in one of two opposite outcomes (Mertler & Vannatta, 2005). Since it does not make sense to regress a continuous independent variable on a binary outcome variable, the LR transforms the data into probabilities that calculate odds ratios that a given value will result in one of two outcomes (Field, 2009). Probabilities are calculated for the outcome given the known values of X1, X2….Xs. Thus, the probability that each occurring case belongs in a given category can be computed. LR is a form of model estimation where MLE is used to select the coefficients that best estimate the known outcome of the categorical variables. The measure used is the log likelihood which is an indicator of the unexplained variance after predictor, and actual outcomes are calculated. The larger the log likelihood, the more cases do not fit the model.

In this research study, the latent variables resulting from the factor analysis were used to predict membership in two opposite groups: participants who were rated with significant cognitive executive dysfunction and those who were not. Membership in the ED group was established by a significant parent rating on the Metacognition Index of the BRIEF. A significant rating will be defined in terms of the T-scores used on the BRIEF. T-scores at least two standard deviations above the mean (T-score > 65) were used for classification in the ED group. To avoid errors in classification, cases included in the NOED group were comprised of participants with scores at or below one standard deviation above the mean (T < 60).

Use of latent variables in Logistic Regression. To use latent variables in LR, different methods of data transformation were used. These methods are classified as unrefined and refined methods (Distefano & Mîndril, 2009) . Unrefined methods involve the averaging of scores in each case based on their factor weights. The drawback to such transformations, which are common in research, is that they do not create a set mean and standard deviation for the factor scores. In addition, the correlations between factors may not be accurately represented. Refined methods of data transformation are more complex and more accurately represent the factor structures. Of the available methods of saving datasets that match the factor structure, the regression method was deemed most appropriate. This method maximizes validity in further analysis. It takes into account the correlation between factors and observed variables through item loadings. It also takes into consideration the correlation among observed variables and between factors. Further, it is most appropriate for factor analysis conducted with oblique rotation. The datasets for the LR were derived using this method by assigning hypothetical factor scores to each case in the dataset. The resulting scores are interval scores similar to z-scores with values varying between -3.0 and +3.0.

Model testing. To measure model fit, a number of statistics were generated in the LR. The first step was to observe the percentage of correctly classified cases. Simply put, could scores on the latent variables predict the group in which a student belonged. Higher scores on tasks should result in a greater chance of lower observe scores on the BRIEF Metacognitive Index. The LR then generates a list of the unstandardized and standardized beta coefficients for each variable in the model (Mertler & Vannatta, 2005). The beta coefficients represent the individual contributions to the model of the latent variables. A significance test for each variable known as the Wald statistic assesses the significance of each variable in its contribution to the model, which indicates which latent variables best explains differences in the outcome groups.

The Hosmer and Lemeshow test is a Chi-Square that assesses how well the data fits the predictive model. If the test is non-significant, the data generally fits the proposed model. The significance of the model is also assessed using a Chi-Square analysis, which compares the appropriateness of the resulting model to a fixed (constant) model. If the two are significantly different ($p < .05$), the latent variables will have been shown to predict group membership beyond chance. A Nagelkerke squared correlation index can then be used to indicate the amount of variance in the model accounted for by the variables. The greater the effect size, the greater the influence of the researched variables on the outcome.

The final statistic of interest is the odds ratio represented by Exp (*B*). The Exp (*B*) expresses the expected change in classification based upon a unit change in one predictor variable. In respect to the current research, an increase in scores in the predictor variables should also reduce the odds of being classified in the ED group. The odds ratio can used to determine the extent to which the value of the latent variables would need to be reduced or increased to significantly reduce or eliminate false positives (students wrongly classified as impaired) and false negatives (students wrongly classified as not impaired). If the Exp (*B*) is sufficiently large, odds ratios can be computed.

Analysis of acceptable sample size. The sample size for logistic regression is dependent on the number of predictor variables used. This is one reason variables were combined in the EFA. With an expected three to four predictor variables, sample size in logistic regression should be adequate at the 10 events per predictor variable rule of thumb, especially in light of research showing that the likelihood of a type I error is low (Vittinghoff & McCulloch, 2007). Type II error is greater with low events per variable. Small sample sizes generally show robust results if there is balance of cases in the outcome groups and response bias in measures is not excessive. To ascertain necessary sample size, G*Power software was utilized (Faul, Erdfelder, Lang, & Buchner, 2007). Assuming ample statistical power ($1-\beta = .80$), a sample of 64 would be considered adequate for detecting effect sizes above .10. Therefore, the sample used in the EFA was deemed to be more than adequate to conduct an LR analysis. Logistic regression analysis is robust in its ability to violate data assumptions of other statistical methods. It is highly susceptible to outliers which can be tested and removed prior to the analysis.

CHAPTER IV

RESULTS

This chapter presents the results of the statistical analyses as outlined in the previous section. The purpose of these analyses was to explore the factor structure of assessment instruments evidencing empirical validity in the measurement of executive processes. Once a valid factor structure was obtained, the latent variables were used to discern if these instruments predicted observer ratings of executive competency or impairment.

Preliminary Analysis

Initially, 204 cases were extracted from the larger database that contained parent/guardian ratings on the BRIEF Metacognitive Index and met the age limits for the study. In cases where more than one parent completed the BRIEF, the mean of the two scores was used. Eight cases were excluded due to BRIEF T-scores falling outside the ranges specified in Table 2. Cases were then analyzed for indicators of overall cognitive ability. Since overall cognitive ability scores were not reported for each case, a secondary analysis was conducted using core verbal and fluid reasoning measures of intelligence. If either overall cognitive ability or indicators of both verbal and fluid reasoning ability fell below one standard deviation from the mean, the case was excluded. For cases where core measures of reasoning ability were not readily available, core academic scores were examined (math reasoning, reading comprehension, written expression, oral language). If these scores were all below one standard deviation from the mean, the case was excluded. Using these criteria, another 10 cases were eliminated. A missing values analysis using SPSS was then completed to eliminate cases reporting less than 5 study variables. Another 7 cases were eliminated, leaving 179 cases for analysis.

Frequency statistics were completed to assess the percent of each variable missing from the dataset. The NEPSY-II Inhibition: Switching and Word List Interference tasks were eliminated because they did not meet the frequency (≥ 45% of cases with the variable present) for imputation. Variables considered highly related to one another were also examined for collinearality through linear regression analysis (e.g. D-KEFS Color Word, Condition 3 and Condition 4, the NEPSY-II Auditory Attention and Response Set) but no extreme correlations were found. Therefore, of the initial 21 variables, 19 were selected for data imputation. Ten datasets were imputed, constraining results to the range of scores found within the sample. Imputation converged in 50 iterations or less. The datasets were then pooled into a single dataset representative of the original data. A regression analysis was conducted comparing the set of means from the original dataset with the means obtained from the imputed dataset. This process was repeated for the standard error of the mean and standard deviations of the original and imputed datasets. The correlation between original and imputed means was extremely high (r = .99), as were the correlations between the standard mean error values (r = .98) and standard deviation values (r = .98). In addition, an independent samples t-test of each variable revealed no statistically significant differences between original and imputed values.

Statistical Assumptions

Assumptions of normality were assessed. Using the Kolgomorov-Smirnov test of normality, eight variables were found to be significant for non-normal distribution. Further evaluation of skewness and kurtosis values indicated all distributions fell within the +/- 1 tolerance range for skewness, and the +/- 3 value for kurtosis. The deviation from normality was therefore judged too small to violate assumptions of normality. Each case was examined using the Mahanoblois procedure to ensure no extreme values existed. None of the 179 cases was statistically significant for outliers.

Descriptive Statistics

The final sample consisted of 179 cases containing 119 children between 8 and 12 years of age and 60 between the ages of 12 years, 1 month and 16 years of age. The sample was predominately comprised of male participants (68%). Given the higher frequency of clinical diagnoses among school aged males (Gershon, 2002), a higher representation of males was expected in the sample. Ethnicity was reported for 153 of the 179 cases. Of these, 109 reported as Caucasian, 17 reported as African American, 22 reported as Hispanic/Latino, and 5 reported as “Other.” Crosstabulation analysis using Pearson’s Chi-Square was conducted to examine the relationship between participants’ gender, ethnicity and age. There were no significant relationships between age and ethnicity χ^2 (5) =5.99, p = *ns,* age and gender, χ^2 (1) =1.36, p = *ns* or ethnicity and gender χ^2 (5) =6.82, p = *ns*.

Descriptive statistics for the complete dataset are presented in Table 3, including previously discussed measures of skewness and kurtosis. Means and standard deviations of variables separated by gender and ethnicity are presented later in this section. Overall, means and standard deviations for the subtests were generally lower than those in the standardization populations for the tests. This result was expected given the sample was drawn from a clinical population. Means for the WJIII COG NU Visual Matching, the D-KEFS Trail Making, Condition 4, and the NEPSY-II Inhibition: Inhibition had the lowest mean scores in the sample, close to one standard deviation below the normative mean. Means for the D-KEFS Tower, and the WJIII COG NU Concept Formation and Analysis Synthesis were closest to the normative sample test means. The standard deviations of each subtest were lower than those in the normative sample indicating less variability then would be expected in a mixed clinical sample. The mean score of the BRIEF Metacognitive Index was elevated above the normative sample mean by approximately 1.5 standard deviations, indicating a majority of participants in the clinical sample were rated as having above average impairments in the subscales comprising the Metacognitive Index.

Table 3

Descriptive Statistics for the Full Imputed Sample

Variable	M	SE	SD	Kurtosis	Skewness	Min	Max
WJIII COG NU							
CANC	95.82	.78	10.46	.31	.06	67.00	125.00
VISM	85.73	.99	13.23	.51	-.02	51.00	126.00
DSPD	92.22	1.07	14.29	.03	.21	60.00	133.00
NR	91.75	1.04	13.94	.34	-.38	49.00	127.00
AWM	93.29	1.04	13.97	.47	-.24	50.00	127.00
CFRM	99.16	1.08	14.52	.36	-.18	49.00	139.00
ASYN	99.71	.90	12.02	.65	-.47	58.00	125.00
D-KEFS							
TMT4	7.25	.26	3.49	-.66	-.10	1.00	16.00
VF3	8.88	.24	3.23	.31	-.12	1.00	19.00
DF3	8.54	.19	2.60	.76	.25	3.00	18.00
CW3	8.60	.24	3.27	-.22	-.39	1.00	15.00
CW4	8.25	.23	3.04	-.17	-.42	1.00	14.00
20ACH	8.93	.22	2.89	.42	-.48	1.00	16.00
TOW	9.99	.19	2.53	.73	.28	3.00	19.00
CS	9.04	.21	2.76	.20	-.08	1.00	18.00
NEPSY II							
ACS	8.71	.27	3.65	-.03	.27	1.00	19.00
AA	8.78	.24	3.14	-.39	-.40	1.00	16.00
RS	6.97	.26	3.50	-.81	.24	1.00	16.00
IN2	7.30	.23	3.08	-.51	-.09	1.00	14.00
BRIEF							
META	64.13	.95	12.77	-.71	-.09	35.00	94.00

Note. CANC- Pair Cancellation; VISM- Visual Matching; DSPD - Decision Speed; NR - Numbers Reversed; AWM - Auditory Working Memory; CFRM - Concept Formation; ASYN - Analysis Synthesis;TMT4 - Trail Making Test, Condition 4; VF3 - Verbal Fluency Test, Condition 3; DF3 - Design Fluency Test,Condition 3; CW3- Color Word Interference, Condition 3; CW4 -Color Word Interference, Condition 4; 20Q - Twenty Questions; TOW - Tower Test; CS- Card Sorting; ACS – Animal Card Sorting, AA - Auditory Attention; RSP - Response Set; IN2 - Inhibition: Inhibition, Metacognitive Index

Bivariate Correlations

Bivariate correlations were reviewed for variables that correlated poorly with the other variables in the dataset. Correlations were calculated for the subtests of each assessment battery and with tasks in other batteries.

Table 4 presents the correlations amongst subtests of the D-KEFS. Asterisks indicate statistically significant relationships. In addition, correlations that equal or exceed .30 are italicized. Correlations at or above .30 indicate the variables share moderate correlations; those below .30 indicate weak correlations (Gertsman, 2015). The higher the correlation between variables, the more likely they will form valid factors.

An analysis of the D-KEFS revealed 11 correlations between the tasks that were significant at the .01 level, and six that were significant at the .05 level. Of these 17 significant correlations, 7 tasks correlated at a moderate level or higher. The D-KEFS Trail Making shared the most significant correlations with other D-KEFS tasks and exhibited moderate correlations with both conditions of Color Word Interference as well as Verbal Fluency. Design Fluency exhibited only one significant correlation at the .05 level. The lack of substantial correlations between the empirical categories of D-KEFS tasks suggest that they are most likely measuring different applications of executive skills needed to regulate the cognitive modality of the task being administered.

Table 4

Pearson's Product Moment Correlations within the Delis-Kaplan Executive Function System (D-KEFS) Subscales

	TMT4	VF3	DF3	CW3	CW4	CS	20Q
VF3	*.380* **						
DF3	.100	-.060					
CW3	*.410* **	.250 **	.010				
CW4	*.410* **	.120	.000	*.440* **			
CS	.290 **	*.360* **	.150 *	.170 *	.260 *		
20Q	.170 *	.190 *	.090	-.140	-.110	*.360* **	
TOW	.170 *	.200 **	.050	.020	.070	.230 **	*.300* **

Note. TMT4 - Trail Making Test, Condition 4; VF3 - Verbal Fluency Test, Condition 3; DF3 - Design Fluency Test, Condition 3; CW3- Color Word Interference, Condition 3; CW4 -Color Word Interference, Condition 4; CS - Card Sorting; 20Q - Twenty Questions; TOW - Tower Test
Correlations ≥ .30 that indicate favorable factorability of variable
*p<.05. **p<.01

Table 5 presents correlations within, as well as between, the subtests of the WJIII COG NU and NEPSY-II. Owing to the strong construct validity of the broad-band cognitive abilities of the WJIII COG NU, moderate to strong correlations were found between tasks measuring common constructs. Decision Speed, Visual Matching and Cancellation (broad ability = processing speed) all shared moderate to strong correlations with one another. Concept Formation and Analysis Synthesis (CHC broad ability = fluid reasoning) were moderately correlated, as were Numbers Reversed and Auditory Working Memory (CHC broad ability = short-term memory).

Table 5

Pearson's Product Moment Correlations within/between the WJIII COG NU and the NEPSY-II

	WJIII COG NU						NEPSY-II			
	CANC	VISM	DSPD	NR	AWM	CFRM	ASYN	AA	RSP	ACS
WJ III COG										
VISM	_.610_ **									
DSPD	_.580_ **	_.590_ **								
NR	.220 **	_.300_ **	.170 *							
AWM	.200 **	.110	.040	_.430_ **						
CFRM	.140	.240 **	.180 *	_.420_ **	_.300_ **					
ASYN	.020	.230 **	.200 **	_.390_ **	.200 **	_.490_ **				
NEPSY-II										
AA	.060	.190 *	.060	.170 *	.070	.240 **	.280 **			
RSP	.270 **	.120	.100	_.330_ **	.270 **	.260 **	.100	.240 **		
ACS	.160 *	.170 *	_.340_ **	_.300_ **	.250 **	_.590_ **	_.330_ **	.090	_.320_ **	
IN2	_.480_ **	_.440_ **	_.310_ **	.240 **	.220 **	_.340_ **	.160 *	.240 **	.170 *	_.360_ **

Note. CANC- Pair Cancellation; VISM- Visual Matching; DSPD - Decision Speed; NR - Numbers Reversed; AWM - Auditory Working Memory CFRM - Concept Formation; ASYN - Analysis Synthesis; AA - Auditory Attention; RSP - Response Set; ACS - Animal Card Sort IN2 - Inhibition: Inhibition

Correlations ≥ .30 that indicate favorable factorability of variable are italicized

*p<.05. **p<.01

Memory tasks correlated just as strongly with measures of fluid reasoning, meaning these two constructs may be indistinguishable in the current sample. Most of the tasks shared significant but weak correlations between broad ability categories regardless of the construct being measured. The NEPSY-II tasks all exhibited significant correlations with one another except that Auditory Attention and Animal Card Sorting were virtually unrelated. Auditory Attention exhibited expected significant correlations at the .01 level with Response Set and Inhibition: Inhibition; however, Response Set showed only a weak relationship with the Inhibition: Inhibition task. Even though these tasks were only weakly related, they both exhibited moderate correlations with Animal Card Sort.

Of the 28 correlations between the WJ- III COG NU and the NEPSY-II, 18 were significant at the .01 level and 5 were significant at the .05 level. Of these, nine correlations were moderate. Animal Card Sort and Inhibition: Inhibition accounted for eight of these moderate correlations. Animal Card Sort correlated most strongly with Concept Formation and correlated moderately with Numbers Reversed, Analysis Synthesis, and Decision Speed. Inhibition: Inhibition correlated moderately with all WJ III COG NU processing speed tasks and with Concept Formation.

Table 6 presents the bivariate correlations between the tests of the D-KEFS and the NEPSY-II and WJIII COG NU. Of the 99 correlations, 40 were significant at the .01 level and 14 were significant at the .05 level. There were 23 moderate correlations between the tasks. Of these, the D-KEFS Trail Making had the most significant correlations at the .01 level with other tasks, followed by the D-KEFS Verbal Fluency and Card Sort, which each had seven. Trail Making shared the most moderate correlations with other tasks with eight followed by Card Sort with four. Verbal Fluency shared only two moderate correlations, as did most other tasks. with the exception of Twenty Questions, which shared three moderate correlations. NEPSY-II Auditory Attention and D-KEFS Design Fluency were the least correlated with other tasks and did not share any moderate correlations. The NEPY-II Response Set also failed to share any moderate correlations with tasks of the D-KEFS or WJ III COG NU.

Relationships Between the Measured Subscales and the Demographic Variables

In order to determine the multivariate differences of the subtests from the WJIII COG NU, NEPSY-II and D-KEFS, a one-way ANOVA was calculated for each variable first by gender, age, and reported ethnicity. Table 7 presents the means and standard deviations of the sample by gender. Differences in variable means with regard to gender were non-significant except for the WJIII COG NU Decision Speed task (F (1,177) = 11.02, p =.001), in which females scored significantly higher (M = 97.20, SD = 14.22) than males (M = 89.83, SD= 13.75).

Table 8 presents the means and standard deviations of the sample by age group. Subtest score differences between children 8 to 12 years of age and children over the age of 12 were non-significant; however, a significant difference in the BRIEF Metacognitive Index score emerged (F (1, 177) = 10.39, p = .002) with parents reporting more executive processing difficulties in older children (M = 68.30, SD =12.29) than in younger children

(M = 61.97, SD = 12.52). Thus, though assessment instruments did not detect differences in EF between age groups, observers reported a higher degree of impairment in executive processes in older children than in younger children.

Sample differences were also analyzed in terms of reported ethnicity. Ethnic groups comprising over 5% of the sample were collapsed into three groups: Caucasian, African American, and Hispanic/Latino. Ethnicity was reported in 147 cases and results are presented in Table 9. Differences in subtest scores were significant for the WJIII COG NU Concept Formation (p = .009), D-KEFS Twenty Questions (p =.001), and NEPSY-II Animal Sort (p = .002) tasks. Post Hoc Bonferroni analyses revealed no significant differences between groups for Concept Formation; however, Caucasians scored significantly higher than the two minority groups on the D-KEFS Twenty Questions, and significantly higher than African Americans on the NEPSY-II Animal Card Sort. No other significant differences between groups were found.

Results of the Exploratory Factor Analysis (EFA)

Prior to the EFA analysis, the bivariate correlations were analyzed for each variable to determine their factorability. Two variables, the NEPSY-II Auditory Attention and the D-KEFS Design Fluency, were minimally related to the other variables in the dataset and therefore were eliminated from the analysis. An initial factor analysis using the MLA method of factor extraction was conducted. The resulting determinant was <.00001, and the Kaiser-Meyer-Olkin Measure of Sampling Adequacy was poor (.545).

Table 6

Pearson's Product Moment Correlations between the D-KEFS, WJIII COG NU and the NEPSY-II

	WJIII COG NU							NEPSY-II			
D-KEFS	CANC	VISM	DSPD	NR	AWM	CFRM	ASYN	AA	RSP	ACS	IN2
TMT 4	*.430* **	*.580* **	*.430* **	*.370* **	.190 **	*.350* **	*.410* **	.220 **	.200 **	*.330* **	*.320* **
VF3	.270 **	.210 **	*.340* **	.150 *	.210 **	.050	.170 *	.110	.250 **	*.560* **	.200 **
DF3	.090	.160 *	.020	.230 **	.170 *	.290 **	-.160 *	.250 **	-.060	.090	.130 *
CW3	*.370* **	*.360* **	-.100	.200 **	.160 *	-.060	.180 *	.010	.080	-.070	.030
CW4	.260 **	.300 **	.140	.180 *	.170 *	.170 *	*.370* **	-.090	.070	.080	.250 **
CS	.250 **	.130	.170	*.420* **	*.390* **	*.520* **	.280 **	.130	.250 **	*.670* **	.190 *
20Q	.090	-.050	.190	*.390* **	.210 **	*.430* **	.220 **	.180	.240 **	*.470* **	.180 *
TOW	*.460* **	.280 **	.290 **	.290 **	*.520* **	.110	.010	.190	.190 *	.070	.220 **

Note. TMT4 - Trail Making Test, Condition 4; VF3 - Verbal Fluency Test, Condition 3; DF3 - Design Fluency Test, Condition 3; CW3- Color Word Interference, Condition 3; CW4 -Color Word Interference, Condition 4; CS - Card, Sorting; 20Q - Twenty Questions; TOW - Tower Test; CANC- Pair Cancellation; VISM- Visual Matching; DSPD - Decision Speed; NR - Numbers Reversed; AWM - Auditory Working Memory; CFRM - Concept Formation; ASYN - Analysis Synthesis; AA - Auditory Attention; RSP - Response Set; ACS - Animal Card Sort; IN2 - Inhibition: Inhibition
*$p<.05$. **$p<.01$
Correlations ≥ .30 that indicate favorable factorability of variable are italicized

Table 7

Means and Standard Deviations for Woodcock Johnson III Tests of Cognitive Abilities (WJIIII COG NU), NEPSY-II, and Delis-Kaplan Executive Function System (D-KEFS) Subtest Scores by Child's Gender

	Male		Female				
	Mean	SD	Mean	SD	*F*	*P*	
WJIII COG							
CANC	94.89	10.59	97.77	9.99	3.00	.085	
VISM	84.63	13.93	88.02	11.40	2.59	.109	
DSPD	89.83	13.75	97.20	14.22	11.02	.001	*
NR	92.05	13.95	91.11	14.00	.18	.673	
AWM	93.51	13.74	92.83	14.55	.09	.761	
CFRM	99.09	13.95	99.30	15.76	.01	.930	
ASYN	100.28	12.62	98.52	10.65	.85	.359	
D-KEFS							
TMT4	6.98	3.51	7.81	3.41	2.27	.133	
VF3	8.56	3.38	9.56	2.80	3.79	.053	
DF3	8.48	2.53	8.67	2.75	.22	.636	
CW3	8.63	3.32	8.53	3.20	.04	.851	
CW4	8.21	3.14	8.34	2.86	.08	.779	
CS	9.13	2.76	8.85	2.77	.39	.533	
20QU	8.73	3.05	9.34	2.52	1.74	.189	
TOW	9.83	2.52	10.35	2.55	1.66	.199	
NEPSY-II							
AA	8.85	3.10	8.64	3.25	.17	.678	
RSP	6.91	3.43	7.10	3.67	.11	.735	
ACS	8.48	3.68	9.19	3.59	1.47	.228	
IN2	7.29	3.26	7.32	2.69	.00	.949	
BRIEF							
META	63.91	12.16	64.57	14.05	.10	.749	

Note. CANC- Pair Cancellation; VISM- Visual Matching; DSPD - Decision Speed; NR - Numbers Reversed ; AWM - Auditory Working Memory; CFRM - Concept Formation; ASYN - Analysis Synthesis; TMT4 - Trail Making Test, Condition 4; VF3 - Verbal Fluency; Test, Condition 3; DF3 - Design Fluency Test, Condition 3; CW3- Color Word Interference; Condition 3; CW4 -Color Word Interference, Condition 4; CS - Card, Sorting; 20Q - Twenty Questions, Achievement, TOW-Tower; , AA- Auditory Attention; RSP - Response Set; ACS - Animal Card Sort; IN2 - Inhibition: Inhibition; META - BRIEF Metacognition Index

* Significant at the .05level.

Table 8

Means and Standard Deviations for Woodcock Johnson III Tests of Cognitive Abilities (WJIIII COG NU), NEPSY-II, and Delis-Kaplan Executive Function System (D-KEFS) Subtest Scores by Child's Age Group

	Age 8 to 12		Age 12.1 to 16				
	Mean	SD	Mean	SD	*F*	*P*	
WJIII COG							
CANC	96.43	9.28	94.65	12.42	1.17	.281	
VISM	85.80	13.63	85.59	12.53	.01	.920	
DSPD	93.14	14.25	90.43	14.31	1.46	.229	
NR	91.59	14.19	92.06	13.53	.05	.832	
AWM	94.22	13.62	91.50	14.58	1.53	.218	
CFRM	99.61	14.33	98.28	14.95	.34	.560	
ASYN	100.09	11.53	98.98	12.97	.35	.557	
D-KEFS							
TMT4	7.34	3.70	7.06	3.07	.26	.613	
VF3	9.08	3.12	8.51	3.43	1.27	.261	
DF3	8.32	2.70	8.97	2.35	2.59	.109	
CW3	8.63	3.53	8.54	2.74	.03	.873	
CW4	8.56	3.16	7.65	2.74	3.65	.058	
CS	9.17	2.89	8.79	2.50	.76	.385	
20Q	8.95	2.96	8.88	2.78	.03	.868	
TOW	10.02	2.60	9.95	2.42	.03	.863	
NEPSY-II							
AA	8.60	3.20	9.12	3.02	1.09	.298	
RSP	6.96	3.67	6.99	3.17	.00	.947	
ASC	8.99	3.59	8.17	3.75	2.05	.154	
IN2	7.53	3.05	6.85	3.11	1.98	.161	
BRIEF							
META	61.97	12.52	68.30	12.29	10.39	.002	*

Note. CANC- Pair Cancellation; VISM- Visual Matching; DSPD - Decision Speed; NR - Numbers Reversed AWM - Auditory Working Memory; CFRM - Concept Formation; ASYN - Analysis Synthesis; TMT4 - Trail Making Test, Condition 4; VF3 - Verbal Fluency Test, Condition 3; DF3 - Design Fluency Test, Condition 3 CW3- Color Word Interference; Condition 3; CW4 -Color Word Interference, Condition 4; CS - Card, Sorting 20Q - Twenty Questions; TOW - Tower Test; AA - Auditory Attention; RSP - Response Set; ACS - Animal Card Sort; IN2 - Inhibition: Inhibition; META - BRIEF Metacognition Index

* Significant at the .05 level

Table 9

Means and Standard Deviations for Woodcock Johnson IIP Tests of Cognitive Abilities (WJIIII COG NU), NEPSY-II, and Delis-Kaplan Executive Function System (D-KEFS) Subtest Scores by Ethnicity

	Caucasian		African American		Hispanic/Latino				
	Mean	SD	Mean	SD	Mean	SD	*F*	*P*	
WJIIII COG									
CANC	94.87	9.78	94.76	9.87	98.33	12.76	1.07	.345	
VISM	85.20	13.90	86.55	12.29	85.25	14.17	.07	.931	
DSPD	92.89	14.35	88.19	16.44	92.47	13.19	.78	.459	
NR	91.50	13.66	93.27	15.71	87.99	13.45	1.00	.450	
AWM	92.85	14.96	95.45	12.26	93.42	13.49	.93	.787	
CFRM	101.63	15.45	92.61	14.09	93.30	11.20	.80	.009	*
ASYN	100.88	11.72	96.79	8.05	96.00	15.65	1.53	.132	
D-KEFS									
TMT4	7.40	3.55	5.81	3.57	7.26	3.25	1.30	.221	
VF3	9.01	3.25	8.55	3.46	7.82	2.97	.49	.277	
DF3	8.65	2.61	7.97	2.23	8.67	3.14	.04	.612	
CW3	8.29	3.49	8.49	1.80	8.42	3.00	.39	.964	
CW4	8.15	2.90	7.46	3.89	8.02	2.88	.24	.678	
CS	9.27	2.91	8.00	2.44	8.10	2.57	2.69	.071	
20Q	9.59	2.63	7.74	2.66	7.42	3.28	6.66	.001	*
TOW	9.96	2.41	10.63	3.40	10.30	2.06	8.01	.538	
NEPSY-II									
AA	8.98	3.37	7.82	2.83	9.06	2.72	.62	.371	
RSP	6.76	3.61	7.55	3.33	6.02	2.95	4.93	.399	
ASC	9.42	3.77	7.93	3.64	6.44	2.97	2.05	.002	*
IN2	7.37	3.15	6.73	3.35	7.29	2.67	1.87	.735	
BRIEF									
META	65.58	12.50	59.59	11.85	62.82	13.94	.31	.157	

Note. CANC- Pair Cancellation; VISM- Visual Matching; DSPD - Decision Speed; NR - Numbers Reversed AWM - Auditory Working Memory, CFRM - Concept Formation; ASYN - Analysis Synthesis; TMT4 - Trail Making Test, Condition 4; VF3 - Verbal Fluency Test, Condition 3; DF3 - Design Fluency Test, Condition 3 CW3- Color Word Interference; Condition 3; CW4 -Color Word Interference, Condition 4; CS - Card, Sorting 20Q - Twenty Questions; TOW - Tower Test; AA - Auditory Attention; RSP - Response Set; ACS - Animal Card Sort; IN2 - Inhibition: Inhibition; META - BRIEF Metacognition Index
* Significant at the .05 level

An analysis of the anti-aliasing matrix was conducted for variables that did not account for an adequate amount of the variance in the factored model. Values below .50 indicate a variable did not have adequate sampling to be included in the analysis. The DKEFS Color Word Interference Task Condition 3 did not meet criteria for sampling adequacy after extraction (MSA = .352) and was excluded from further analysis. Removing this variable resulted in acceptable sample adequacy (Kaiser-Meyer-Olkin = .73) and a Determinant of .001. A Heywood case occurred in the analysis allowing for the NEPSY-II Animal Card Sort task to have a factor loading >1.0. The ULS method of factor extraction was attempted, but did not alleviate the Heywood case. Linear regression analysis ruled out possible combinations of variables that were collinear with the NEPSY-II Animal Card Sort task, a common cause of Heywood cases. Next, the residual correlations from the reproduced correlation matrix were examined to determine if large negative residuals existed between similar tasks. The NEPSY-II Animal Card Sort, WJIII COG NU Concept Formation, D-KEFS Card Sort, and D-KEFS Verbal Fluency tasks all had negative residuals. The generation of the Heywood case was the result of fluctuations in the data that invalidated the mathematical process of MLA.

To alleviate the Heywood case, various solutions were attempted. First, Animal Card Sort was removed from the analysis since it had the factor loading of greater than 1.0; however, this did not alleviate the Heywood case in either MLA or ULS extraction. Next, other negatively correlated variables were removed beginning with the D-KEFS Verbal Fluency task, specifically because it had the least moderate correlations with other variables. This removal alleviated the Heywood case in the ULS analysis but not in the MLA; thus, the analysis was continued using the alternative ULS extraction method.

In addition to the D-KEFS Verbal Fluency being omitted, the NEPSY-II Response Set was also omitted due to an exceptionally low extracted communality. In the final analysis, this task accounted for less than 3% of the variance in the model. The analysis was generated with the remaining 14 variables using direct oblique rotation due to anticipated correlations among the generated factors. Measures of validity were analyzed first. The Determinant of the correlation matrix was within tolerance levels at .002. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was adequate at a value of .74. KMO and Bartlett's Test of Sphericity indicated the matrix was not an identity matrix (χ^2 (91) = 1053.58, p <.0001). The PCA Parallel Analysis simulation indicated that three factors emerged at a level greater than chance. An analysis of the screen plot found that three or four factors would yield a viable solution. Therefore, both models were explored. The data was first analyzed with four factors extracted accounting for 67.2% of the variance among the variables. Factor loadings are presented in Table 10. To facilitate interpretation, factor loadings below .30 were suppressed in the output and do not appear in the table. In terms of best practices for interpretation, a loading of .40 indicates adequate reliability for interpretation of a factor loading. Values above .30 indicate a moderate relationship between the variable and that factor. Cross loadings above .32 indicate a complex variable that contributes to more than construct.

Table 10

Factor Structure of the Four Factor Model

Subtest	Factor Description (Reliability)				Total Communality
	Cognitive Flexibility (.70)	Selective Attention (.86)	Fluid Reasoning (.89)	Working Memory/Planning (.79)	
WJCANC		**.77**			.70
WJVISM		**.81**			.73
WJDSPD		**.71**			.55
WJNR	.33			.35	.43
WJAWM				**.67**	.55
WJCF			**.61**		.59
WJAS	**.58**				.49
DKCS			**.55**		.50
DKTMT	**.40**	**.49**			.55
DKCW4	**.56**				.38
DK20Q			**.61**		.43
DKTOW				**.79**	.72
NP2ACS			**.93**		.84
NP2IN2		**.43**			.32

Note. WJCANC- Pair Cancellation; WJVISM- Visual Matching -WJDSPD - Decision Speed; WJNR - Numbers Reversed; WJAWM - Auditory Working Memory; WJCFRM - Concept Formation; WJASYN - Analysis Synthesis; DKTMT4 – Trail Making Test, Condition 4; DKCW4 -Color Word Interference, Condition 4; DKCS - Card, Sorting; DK20Q - Twenty Questions; DKTOW - Tower Test; NP2ACS - Animal Card Sort; NP2IN2 - Inhibition: Inhibition

Factor 1 accounted for 33.8 % of the variance in the model (Eigenvalue = 4.73), Factor 2 accounted for 14.6% of the model (Eigenvalue = 2.05), Factor 3 accounted for 10.3% variance (Eigenvalue = 1.44) and Factor 4 accounted for 8.5% variance (Eigenvalue = 1.19). Total extracted communalities all fell above .30 indicating all variables accounted for at least 9% of the total variance. The NEPSY-II Animal Card Sort had the highest overall communality (.84). The D-KEFS Tower, and WJIII COG NU Visual Matching and Cancellation tasks all had communalities exceeding .70. The D-KEFS Color Word Interference, Condition 4 and NEPSY-II Inhibition: Inhibition were the only variables with communalities below .40.

Analysis of the Four Factor Model

The first factor that emerged had three variables loading over .40. Two of these, WJIII Analysis Synthesis and D-KEFS Color Word Interference were strongly associated with Factor 1. The D-KEFS Trail Making loaded adequately on this factor, but had a loading > .40 on Factor 2. The WJIII COG NU Numbers Reversed was associated with Factor 1, but its loading of .33 fell below the .40 minimum. The primary skills assessed within this factor were attentional switching and deductive reasoning. Factor 2 had strong loadings from all processing speed tasks of the WJIII COG NU (> .70). Adequate factor loadings were also present from the NEPSY-II Inhibition: Inhibition, as well as the D-KEFS Trail Making, Condition 4. This factor clearly represented a combination of processing speed and response inhibition best described as Selective Attention. Factor 3 was composed of tasks associated with concept formation and fluid reasoning including the WJIII COG NU Concept Formation, D-KEFS Card Sorting and Twenty Question tasks, and the NEPSY-II Animal Card Sort, which was highly correlated with this factor loading at .93. Although a Heywood case was avoided in the ULS analysis, it was still apparent from the results that the tasks loading on this factor had significant redundancy. Factor 4 was comprised of two tasks with strong loadings: D-KEFS Tower and WJIII COG NU Auditory Working Memory. There was also a cross loading present with the WJIII COG NU Numbers Reversed but, as in Factor 1, its communality of .35 indicated the loading was not sufficient to include in the factor. However, it is important to note that both tasks of working memory had their highest loadings on this factor. The D-KEFS Tower also requires working memory to visualize, and hold in memory, sequences of moves to plan and implement an effective strategy. Therefore, Factor 4 seems to be associated with tasks of working memory and planning.

The anticipated correlation between factors emerged between Factor 1 and Factor 3, which were moderately correlated ($r = .30$). This was due primarily to moderate correlations on both factors from WJIII COG NU Numbers Reversed, and Auditory Working Memory and from the D-KEFS Trail Making and Card Sort. All three of those tasks also shared moderate to strong correlations with Factor 4.

Analysis of the Three Factor Model

The PCA parallel analysis indicated that a three factor model may be a better fit for the data. Therefore, the ULS analysis was completed with three factors being extracted. The results are presented in Table 11. Three factors accounted for 58% of the variance amongst the original variables. All communalities were above .30 with the exception of D-KEFS Color Word Interference, Condition 4, which fell to a value of .20. However, it continued to load adequately on the extracted factors. The D-KEFS Tower task had the highest communality at .84, while the communality for the NEPSY-II Animal Card Sort dropped to .58 allowing for other similar measures to account for more variance within the three factors. Factor 1 shows a strong factor structure with four of the seven tasks exceeding factor loadings of .60.

Factor 1 was a clear convergence of tasks involving concept formation and working memory and was composed of the D-KEFS Card Sort and Twenty Questions tasks, NEPSY-II Animal Card Sort, and the WJIII COG NU Concept Formation, Analysis Synthesis, Numbers Reversed, and Auditory Working Memory tasks.

Factor 2 also exhibits a strong factor structure, with four of the six factor loadings above .60. This included the WJIII COG NU Visual Matching, Cancellation, and Decision Speed tasks, the D-KEFS Color Word Interference, Condition 4, and Trail Making tasks, and NEPSY-II Inhibition; Inhibition task. All of these assess response inhibition and attentional switching necessary in the application of Selective Attention skills.

Table 11

Factor Structure of the Three factor Model

Subtest	Concept Formation/Memory (.85)	Selective Attention /Inhibition (.87)	Self Monitoring/Planning (.81)	Total Communality
	Factor Description (Reliability)			
WJCANC		**.77**	.33	.67
WJVISM		**.91**		.76
WJDSPD		**.64**		.42
WJNR	**.52**			.38
WJAWM	**.42**		.39	.37
WJCF	**.76**			.61
WJAS	**.49**			.37
DKCS	**.70**			.51
DKTMT4		**.61**		.54
DKCW4		**.40**		.20
DK20Q	**.61**			.38
DKTOW			**.88**	.80
NP2ACS	**.76**			.58
NP2IN2		**.45**		.31

Note. WJCANC- Pair Cancellation; WJVISM- Visual Matching -WJDSPD - Decision Speed; WJNR – Numbers Reversed; WJAWM - Auditory Working Memory; WJ CFRM - Concept Formation; WJASYN - Analysis Synthesis; DKTMT4 – Trail Making Test, Condition 4; DKCW4 -Color Word Interference, Condition 4; DKCS - Card, Sorting; DK20Q - Twenty Questions; DKTOW - Tower Test; NP2ACS - Animal Card Sort; NP2IN2 - Inhibition: Inhibition

Factor 3 was comprised primarily of the D-KEFS Tower task and moderate cross loadings from the WJIII COG NU Cancellation and Auditory Working memory tasks. Although this factor was difficult to interpret, the three tasks comprising it required consistent self-monitoring and focused attention, while also holding a sequence of objects/information in memory to attain a goal. This factor is best labeled as Self-Monitoring/Planning. The primary loading of only one variable on this factor makes the validity of interpretation of a third factor questionable. However, a large amount of variance was extracted from the Tower task, and when the analysis is run without this task, the relationship between the remaining variables changes significantly. The same set of tasks also emerged in both the three and four factor models and therefore, the interaction between the variables in Factor 3 can be viewed as necessary to the overall factor structure between the remaining variables in the study.

Comparison of the Three and Four Factor Solutions

Both models fell slightly below the recommended CFI threshold of .90 (four factor model CFI = .88, three factor model CFI = .81). The Non-Normed Fit Index (NNFI) for the three factor model (NNFI = .66) indicated a slightly poorer fit than the Four factor model (NNFI = .77). The GFI for both models exceeded .98 indicating that the structure of the two models in explaining variance among the variables was acceptable. Both models also met Kelley's Criterion by falling below the .74 threshold for the RMSR.

Given that fit indices are marginally acceptable for both models, a decision had to be made on which factor structure to retain for the subsequent analysis. An EFA is by definition exploratory, meaning that there are no agreed upon inferential statistical procedures that validate a given hypothesis or theory (Costello & Osborne, 2005). Despite fit indices and guidelines for interpretation and retention, there is no theoretical measure of "correctness" for EFA and the results are almost entirely driven by the subjective decision of the researcher providing the rationale for interpretation (Henson, 2006). Although the four factor model has better overall indices of fit and each factor has at least two primary factor loadings, there are reasons to favor the three factor model. The communalities in the three factor model capture only slightly less variance than in the four factor model but in a more parsimonious way. In addition, PCA parallel analysis, considered by many researchers to be the best measure of factor retention (Costello & Osborne, 2005; Henson, 2006; Taherdoost, et al., 2004) indicated extraction of three factors was optimal. These facts aside, the decision of which analysis best represents the data comes down to both practicality and theoretical validity. While the four factor model hints at possible further delineation of factors, the existing data and sample size does not support more than three factors. In addition, there appears to be some variables whose correlation in the sample may be an artifact of the sample scores more than a solid theoretical relationship between the tasks (e.g. WJIII COG NU Analysis Synthesis factoring with tests of Selective Attention/Attentional Switching in the four factor model). The three factor model approximates a clearer delineation of both the theoretical models presented as well as a more valid interpretation of the factors due to a greater number of loadings per factor. It also eliminates some of the correlations between factors in the more complex model. Thus, if the main purpose of the research is to clarify the structure of EF in the sample, the three factor model presents a more practical choice for further analysis.

Age Related Differences in Factor Structure

Due to significant differences in observer levels of executive dysfunction by age group, the analysis was run separately for each group. Factor loadings differed marginally and differences did not exceed .06 for any communality. Despite significant differences in reported functioning, the factor structure was identical in both age groups.

Results of the Logistic Regression (LR) Analysis

The secondary analysis sought to answer whether the latent variables derived from factors in the first analysis can predict observer reported dysfunction in cognitive executive processes such as working memory, organization, planning, and self-monitoring. To complete this analysis, a binary logistic regression was conducted using the latent variables derived from the three factor model of EF. As mentioned previously, factor scores for each participant based on the factors derived from the EFA were computed. These scores represent a weighted value of the assessment scores in each participants profile based upon the variance in the EF factors captured by each subtest. The derived scores used in the analysis are Z-scores. Before conducting the LR, descriptive statistics were computed to see if the latent variable scores differed based on demographic variables in the sample.

Descriptive data is presented in Table 12 based on ethnicity, age, and gender. These scores were converted from Z-scores to standard scores to assist in interpretation and analysis. Scores are presented for gender for the entire sample (n=179) and for the three main reported ethnicity classifications (n=150). To explore possible sample differences on gender and ethnicity, a Multivariate Analysis of Variance (MANOVA) was conducted with the factor variables as the dependent variables and gender and ethnicity as the independent variables. No significant main effects were present for gender and no interaction effects occurred between gender and ethnicity. A significant main effect of ethnicity was found for Factor 1, Concept Formation/Working Memory (F (7.15, 2) =2.51, p =.039, η =.05). The effect size was small. No other differences between ethnic groups were found. Post Hoc Bonferonni analyses revealed a significant difference in scores on Factor 1 between Caucasians and Hispanic/Latinos in the sample (p = .03) but all other group relationships were non-significant. An analysis of the group means revealed the difference occurred mostly between males in the groups. Despite the small effect size of ethnicity, the LR was computed twice; once for the sample as a whole, and once for participants who identified ethnicity as Caucasian (n = 109).

Table 12

Standardized Means and Standard Deviations for Derived Factors by Reported Gender and Ethnicity

		Caucasian			African American			Latino/Hispanic			Total By Ethnicity			Sample Total by Gender		
		N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
CF/WM	Male	76	102.53	13.53	8	95.51	15.56	18	90.89	14.90	102	99.92	14.53	121	99.91	14.51
	Female	33	102.81	18.51	9	93.27	13.63	6	99.23	11.41	48	100.57	17.11	58	100.19	16.23
	Total	109	102.61	15.12	17	94.32	14.14	24	92.97	14.35	150	100.13	15.35	179	100.00	15.04
SA/RI	Male	76	97.52	15.61	8	91.98	14.00	18	100.45	16.37	102	97.60	15.61	121	98.52	15.43
	Female	33	104.65	13.31	9	100.97	18.09	6	99.04	14.75	48	103.26	14.29	58	103.08	13.81
	Total	109	99.68	15.25	17	96.74	16.46	24	100.10	15.68	150	99.41	15.38	179	100.00	15.04
SM/PLAN	Male	76	98.45	14.73	8	100.61	13.75	18	102.40	12.55	102	99.32	14.25	121	99.28	15.20
	Female	33	100.69	13.32	9	109.15	21.47	6	100.28	15.24	48	102.23	15.34	58	101.51	14.73
	Total	109	99.13	14.30	17	105.13	18.23	24	101.87	12.95	150	100.25	14.62	179	100.00	15.04

Note: CF/WM – Concept Formation and Working Memory; SA/RI – Selective Attention, Response Inhibition; SM/PLAN – Self Monitoring/Planning

Additional analyses were not performed on the other ethnic groups in the sample due to insufficient sample size for the LR. The method used to generate the models was the Enter method meaning all three latent variables were simultaneously entered into the model regardless of the strength of their individual contribution. The initial analysis revealed no statistical outliers defined as a score greater than two standard deviations above the mean for any of the three latent variables.

Latent Variable Prediction of Executive Dysfunction

Results of the LR for the entire sample and the Caucasian group are presented in Tables 13 and 14. In the full sample, all factors were non-significant in correctly classifying the ED versus NOED groups. The Hosmer and Lemeshow test of model fit was significant ($\chi2$ (8) = 18.24, p = .019) indicating the model did not fit the data. The Nagelkerke R^2 indicated the factors accounted for less than 3% of the variance in the model. An analysis of actual versus predicted case classification revealed that 91% of cases in the ED group were correctly classified; however, only 17% of cases in the NOED group were correctly identified, for a total correct classification rate of 59%. An analysis of the case classification plot did not reveal any pattern that would indicate that a change in the cut-off points of actual group membership would improve the model fit. Individuals scoring higher on Factor 1 would be 1.14 times more likely to be classified as NOED, and those scoring higher on Factor 2 would be 1.12 times more likely. Factor 3 exhibited a slight inverse relationship with participants scoring one unit higher being .88 times as likely to be classified in the NOED group.

Table 13

Logistic Regression of Executive Function Factors Predicting BRIEF Metacognitive Index Ratings For Full Sample

Independent Variables	*B*	SE	Wald	Sig	Exp (*B*)
Concept Form/Working Mem	.130	.166	.614	.433	1.139
Selective Attention/Inhibition	.193	.166	1.355	.244	1.213
Self -Monitoring/Planning	-.080	.154	.265	.606	.924
Model $\chi 2$	3.211	p = .36			
Nagelkerke R^2	.02				
Homser & Lemeshow	18.24	p = .02*			
N	179				

Table 14

Logistic Regression of Executive Function Factors Predicting BRIEF Metacognitive Index Ratings For Caucasian Group

Independent Variables	*B*	SE	Wald	Sig	Exp (*B*)
Concept Form/Working Mem	.749	.262	8.151	.004*	2.115
Selective Attention/Inhibition	.145	.229	.398	.528	1.156
Self -Monitoring/Planning	-.353	.235	2.256	.130	.702
Model $\chi 2$	14.056	p = .003*			
Nagelkerke R^2	.17				
Homser & Lemeshow	11.77	p = .162			
N	109				

Results of the Logistic Regression for the Caucasian sample indicated the overall model was a significant predictor of group membership ($\chi 2$ (3) = 14.06, p = .003). The analysis revealed that in the Caucasian group, Factor 1, Categorical Thinking, was a significant predictor; however, Factor 2, Selective Attention/Inhibition and Factor 3, Self -Monitoring/Planning did not add to the model. Overall, 66% of cases were correctly classified. The Hosmer and Lemeshow model fit was significant indicating the model adequately fit the data ($\chi 2$ (8) = 11.77, p =.162). The Nagelkerke R2 indicated the independent variables explained 16.5% of the variance in the model. Exp (B) were again small, with a unit rise in Factor 1 resulting in a 1.2 times likelihood of being classified in the NOED group.

Post hoc analysis was done to discover which variables comprising the factor were significant predictors of group classification. An LR was run using the original variables from the sample that loaded on Factor 1. Although the original variables contained variance due to error not included in the factors, prediction rates were assumed similar. The results of the LR are presented in Table 15. The original variables were also a significant predictor of group membership with 87% of the ED and 40% of the NOED group correctly predicted for a total prediction rate of 70% ($\chi 2$ (7) = 18.611, p = .009). This is opposed to a 48% prediction rate in the remaining 58 cases (African American, Hispanic, and unidentified). Overall, 21% of the variance in group membership was explained by the variables in the model. However, of the variables in the model, only the WJIII COG NU Concept Formation was significant in predicting group membership

($\chi2$ (1) = 3.97), p = .046), and accounted for 18% of the variance explained in group membership. The remaining variables added little to the prediction model and percentage variance explained. The WJIII COG NU Numbers Reversed and NEPSY–II Animal Card Sort tasks were the only other variables to add to the overall prediction rate; however, they accounted for less than 3% of the variance explained. Thus, of the latent variables in the three factor model, Factor 1 had a weak but significant effect on group prediction with the WJIII COG NU Concept Formation task accounting for nearly all of the variance explained by the variables comprising the factor.

Table 15

Logistic Regression of Factor 1 Variables Predicting BRIEF Metacognitive Ratings from Original Data

Independent Variables	*B*	SE	Wald	Sig	Exp (*B*)
WJIIICF	.628	.315	3.97	.046*	1.87
WJIII AS	.021	.280	.01	.939	1.02
WJIII NR	.406	.303	1.79	.181	1.50
WJIII AWM	-.332	.244	1.85	.173	.72
D-KEFS 20Q	-.395	.299	1.73	.187	.67
D-KEFS CS	.094	.322	.08	.771	1.10
NEPSY-II ACS	.449	.365	1.51	.219	1.57
Model $\chi2$	18.61		$p = .009$*		
Nagelkerke R^2	.22				
Homser & Lemeshow	6.14		$p = .631$		
N	109				

Summary of Findings

The purpose of this chapter was to summarize the results of the statistical analyses utilized in this research study. The variables in the study that exhibited moderate correlations or higher in the bivariate correlation matrix were analyzed using exploratory factor analysis. Adjustments were made in both selection of the variables and in the factor methodology used due to poor sampling adequacy of one variable and Heywood cases resulting from fluctuations in the values of the variables. Both a three factor and four factor model were obtained with both generally showing an adequate fit to the data. Although the four factor model showed better fit, it did not explain the relationship among the variables as well as in the three factor model. Factor 1 represented categorical reasoning, and cognitive flexibility skills necessary to form concepts as well as working memory skills, Factor 2 represented Selective Attention and underlying skills of response inhibition and set shifting, while Factor 3 represented self-monitoring and planning skills. The first two factors had strong loadings from multiple tasks. The second analysis analyzed the ability of the latent variables in the three factor structure to predict observer ratings of executive dysfunction. The results indicated the latent variables were non-significant predictors of observed executive dysfunction as defined by parent ratings on the BRIEF Metacognitive Index. The problem was in the identification of false positives (test scores predicting executive dysfunction in the absence of significant parent ratings). The model for predictions among participants identifying with a Caucasian ethnicity was significant with the first factor predicting group membership with 70% accuracy and explained 19% of the variance in-group membership. In subsequent analysis, the WJIII COG NU Concept Formation was the only significant predictor variable loading on this factor.

CHAPTER V

DISCUSSION

The current study explored the factor structure of executive functions in a mixed clinical sample of children the Woodcock Johnson Tests of Cognitive Abilities, 3rd Edition, Normative Update (WJIII COG NU; Woodcock et al., 2007) , the NEPSY-II, a Developmental Neuropsychological Assessment (NEPSY-II; Korkman et al., 2007a), and the Delis Kaplan Executive Function System (D-KEFS; Delis et al., 2001b). An exploratory factor analysis (EFA) was used to derive latent variables measuring facets of executive processes. These latent variables were then used to assess whether formal assessment instruments could predict parent ratings of significant executive dysfunction versus subclinical levels of executive dysfunction as represented by scores on the Metacognition Index of the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia et al., 2000b).

Purpose of the Study

The primary question in this research study addressed the factor structure of EF in a clinical sample using common measures for the neuropsychological assessment of school-aged children. Previous research has examined the relationships between the D-KEFS and the WJIII COG NU (Floyd et al., 2010), and the factor structure and relationships among D-KEFS tasks (Latzman & Markon, 2010). In addition, previous research involving data from the KIDS INC. clinical sample was used in a CFA to clarify relationships between the original NEPSY, D-KEFS and WJIII COG (Avirett, 2011). This previous research revealed that the factor structure of EF is anything but consistent and easily discernible. The difficulty arises in the breadth of tasks purported to measure different cognitive and executive constructs (Packwood et al., 2011), as well the degree of task impurity or tasks that assess multiple cognitive and executive skills (Miyake et al., 2000). The current research centered on first discerning similarities and differences in the factor structure of EF in a small clinical sample compared to models evidenced in research. The tasks were chosen because they were common assessment techniques used in school based evaluations and had historically shown empirical validity in assessing EF. The tasks were also chosen to minimize reliance on previous learning and vocabulary, resulting in only one task being included where language was a factor. Some tasks assessed specific skills, while others centered on the ability to solve novel problems such as drawing inferences and developing effective strategies. It was hoped that the emergent factors would approximate current theoretical models in a way that would continue to validate commonly debated constructs. The second purpose was to find congruence between observer ratings of executive difficulties, and results on standardized norm referenced direct assessment of EF. Previous attempts have been made to reconcile direct assessment measures such as the D-KEFS, with rating scales purporting to measure similar constructs such as the BRIEF, but no association has been found (Bishop, 2010; Bodnar, Prahme, Cutting, Denckla, & Mahone, 2007). Since the tasks analyzed and the

BRIEF rating scale are both commonly used in school-based assessments, it was hoped that, in a clinical sample where academic and social achievement deficits were significant, both observer ratings and assessments would discern between average and significant EF.

Summary of Results

Initial analysis of the tasks used revealed little differences in the measures selected according to gender and ethnicity. With few exceptions, these tasks proved to be gender and culturally unbiased within the sample. Females performed significantly better than males on the WJIII COG NU Decision Speed task, and performed marginally better on other tasks of processing speed. Minority groups tended to score slightly lower on most tasks, with the exception of the D-KEFS Tower and Color Word Interference tasks, and significantly lower on the D-KEFS Twenty Questions task, the NEPSY-II Animal Card Sort task, and the WJIII COG NU Concept Formation task. The D-KEFS tasks arguably require the verbal mediation and the WJIII COG NU Concept Formation requires the understanding of complex verbal directions. Both of these factors could have contributed to lower scores in minority samples. The difference could also be an artifact of the much smaller minority representation in the sample.

Overall, participants in the sample scored lower than the normative samples of the three instruments used but still within the average range. Four tasks were an exception to this pattern; the D-KEFS Trail Making, the NEPSY-II Response Set and Inhibition, and the WJIII COG Visual Matching. Interestingly, all of these tasks represented processing speed, response inhibition, and attentional switching with both a verbal and non-verbal response. An analysis of the bivariate correlations indicated that, with the exception of the NEPSY II Response Set, these variables had moderate to strong correlations and all loaded on the second factor in the three factor model.

Further analysis of the bivariate correlations indicated low to moderate correlations within the NEPSY-II and the D-KEFS batteries. Since these instruments do not utilize factor structure or composite factor scores, this is not a surprising result. Previous research on the factor structure of the D-KEFS revealed that the scores from the Card Sorting and Verbal Fluency tasks comprised independent factors while Color Word Interference and Trail Making tasks formed a third factor (Latzman & Markon, 2010). The remaining tasks were unrelated. In the current analysis, D-KEFS Card Sort task shared moderate correlations with the Verbal Fluency and Twenty Questions tasks indicating relationships in the clinical sample that differed from the normative sample. Due to the strong theoretical orientation of the WJIII COG NU, broad ability scores in the areas of processing speed, memory, and fluid reasoning all had moderate to strong correlations. The Numbers Reversed and Auditory Working Memory task also had moderate correlations to fluid reasoning tasks with Numbers Reversed sharing stronger correlations with these tasks than Auditory Working Memory. Previous research has shown a similar association among younger children, ages 9 to 13, but not in older populations (Dombrowski & Watkins, 2013). Scores from the WJIII COG NU processing speed tasks correlated strongly amongst themselves and had moderate to strong correlations with the D-KEFS Trail Making and NEPSY-II Inhibition: Inhibition tasks. However, they had generally weaker correlations with reasoning tasks than these more complex tasks of inhibition and attentional switching. The exception was the WJIII COG NU Cancellation task, which exhibited a moderate correlation with the D-KEFS Tower task.

Overall, the sample scores did not reflect substantial differences in scores compared to the normative sample. Although mean scores were generally lower, they frequently fell within the average range of functioning or slightly below the average range of functioning. This is consistent with previous literature that indicates minor differences in executive processes may have a larger overall effect on performance or, that the use of scaled and standard scores may not provide enough sensitivity to differences in functioning between clinical and non-clinical populations.

Primary Analysis

In the primary analysis, the first research question addressed was:

Is there a discernible factor structure of EF in a clinical sample of children that emerges from commonly used assessment instruments having empirical validity in detecting executive dysfunction? If so, how many factors best describe the data and do these factors represent similar constructs reported in research on adult and child populations?

The primary purpose of this first question was to compare the emergent factor structure of EF tasks to previous literature. The results of the EFA revealed the complexity of EF constructs in child clinical populations with two possible solutions emerging: a four factor model and a three factor model. Both models contained complex variables and exhibited satisfactory reliability of factors. The three factor model was ultimately chosen as the more parsimonious and consistent of the two models in its representation of the data. In this model, two strong factors emerged with the first factor comprising tasks associated with constructs such as concept formation, cognitive flexibility in problem solving, and overall fluid reasoning ability. The WJIII COG NU Concept Formation and Card Sorting tasks from the D-KEFS and NEPSY-II, along with the D-KEFS Twenty Questions, comprised the highest loadings. Weaker loadings occurred with the WJIII COG NU working memory tasks (Auditory Working Memory and Numbers Reversed) and Analysis Synthesis task.

From the correlation matrix, it appeared Analysis Synthesis assessed attentional switching to a degree that its loading was attenuated by its relationship with attentional switching tasks from the D-KEFS (Color Word Interference - Condition 4, and Trail Making, Condition 4). Overall, Factor 1 represents the ability to hold and manipulate information and discard or update that information to generate concepts and draw inferences. Underlying these reasoning tasks are the skills of working memory and cognitive flexibility. This finding adds to the debate as to whether fluid reasoning and working memory are separate and discernible facets of fluid reasoning (Garlick & Sejnowski, 2006). These EFA results provide evidence that, though they are intricately linked in the clinical sample, some differentiation can be made in that the working memory tasks show complex patterns of loadings depending on the number of factors extracted.

Factor 2 was primarily comprised of high loadings from tasks assessing processing speed from the WJIII COG NU. More complex tasks requiring attentional switching between sets from the D-KEFS and NEPSY-II had lower factor loadings, with the verbally mediated inhibition/switching tasks having the lowest factor loadings. In the model extracting four factors, these began to emerge as a separate factor indicating that with a larger sample size, this factor may have delineated between the primary skill of sustained and selective attention for the WJIII COG NU tasks and the more complex skills of attentional switching and response inhibition tapped by the D-KEFS and NEPSY-II tasks. This result is consistent with research indicating that while processing speed, working memory and EF constructs are closely related, EF and working memory, through attentional control, are thought to moderate performance on tasks of processing speed (McCabe et al., 2010). Thus, timed tasks requiring inhibition or switching, though related to processing speed tasks, also exhibited small but significant relationships with tasks of reasoning. Overall, Factor 2 reflected the ability to rapidly process and discriminate information requiring sustained and selective attention with the underlying abilities to inhibit preprogrammed responses and shift between rule sets.

The third factor was the only one to emerge consistently whether three or four factors were extracted, primarily due to only one variable, D-KEFS Tower. As such, it is difficult to interpret this as a valid factor. However, the secondary loadings of both the

WJIII COG NU Auditory Working Memory and Cancellation tasks provide some support for an emergent factor. Performance on the tower task requires holding multiple possible moves in memory to compare if those moves begin to approximate a viable solution. The Auditory Working Memory task also requires that two distinct sets of information be organized, sequenced and retained. Tower tasks have often been investigated as both a task of response inhibition and working memory, with the impact of working memory dependent on the version of the task (Zook, Davalos, Delosh, & David, 2004). For instance, although primarily a visual memory task, the Tower of London task also involves verbal working memory (Phillips, Wynn, Gilhooly, Della Sala, & Logie, 1999). The cancellation task requires both selective and focused attention, in that distracting patterns of images must be ignored to maintain focus. It may also be that the spatial aspect of completing the Tower tasks shares a common skill.

What is clear is that the Tower task resisted convergence with other tasks of fluid reasoning even though working memory skills were important to both factors. This is the same lack of correlation found in the D-KEFS normative sample (Delis et al., 2001b) and other factor analytic research involving the D-KEFS and tasks from the WJIII COG (Floyd et al., 2010). However, Tower's high overall communality and its relationship with the other two measures loading on Factor 3 represented a critical component in maintaining the factor structure of the model. In other words, it could not be removed without confounding the relationships between the other variables. This is evidence that the D-KEFS Tower measures some unique set of skills uncorrelated with other tasks that measure EF. It may be that the most salient skill assessed amongst all three variables was a component of vigilance/self-monitoring requiring resistance to distracting information while performing a sequence of mental or motor steps to reach goal.

Analysis of the EFA Results Compared to Current Theorized Models

The factor structure that emerged was not indicative of the Miyake and Friedman (2000) model. Inhibition did not show a clear pattern of separation whether within a four factor or three factor model. In younger children, this is not uncommon since skills are undifferentiated (Van der Ven, Kroesbergen, Boom, & Leseman, 2013). However, there is evidence that as children age these factors should be more easily discernible (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003). As reported in the results section, the derived factor structure did not differ between pre-adolescent (8 years to 12 years) and adolescent participants (12 years, 1 month and up). This result may indicate that in clinical populations, the factor structure of EF is more stable meaning there is less differentiation and development in the underlying skills. In the current research, the expected differentiation between working memory and set shifting in factors did not emerge. However, it is possible that the self-monitoring component of Factor 3 does reflect a possible convergence of working memory and response inhibition skills.

An argument can be made that the factor structure that emerged is best represented by the factors of CHC theory. The first two factors could easily be named Fluid Reasoning (*Gf*) and Processing Speed (*Gs*). However, this ignores the convergence of working memory and fluid reasoning as well as the presence of the third factor. In fact, Concept Formation exhibited much better correlation and similar factor loadings with the card sorting tasks of the D-KEFS and NEPSY-II than with Analysis Synthesis, and Auditory Working Memory had a higher correlation with the D-KEFS Tower task than it did with the Numbers Reversed task. There is also the previously mentioned difference in factor loadings in Factor 1 between tasks of reasoning and memory, and in Factor 2 between tasks of processing speed and inhibition/shifting. Thus, though CHC factors are represented in the analysis, they do not adequately explain the relationship between variables.

When examining models discussed in the research, the Integrated School Neuropsychological Processing and Cattell-Horn-Carroll model (SNP/CHC) seems to be best reflected by the results. In this model, skills such as working memory and attention are facilitators of integrated processing systems that serve to develop responses too novel problems (Miller, 2013). Among the facilitators of basic cognitive processes are working memory, processing speed and attention. It is clear from the three factor model, the factor comprising the highest degree of variance comprises what the SNP would describe as EF processes of concept formation, problem solving, planning and reasoning, and cognitive set shifting. The additional but substantially lower loadings of working memory tasks, provides evidence that the skills assessed by these tasks are highly related but not central to this factor. In Factor 2, the underlying facilitator skill of processing speed assessed by the Cancellation and Visual Matching tasks comprise the highest factor loadings and can be said to underlie the more complex tasks of inhibition/set shifting. In

Factor 3, secondary loadings of processing speed and working memory tasks may represent the efficiency of underlying facilitation skills that assist in performing the Tower task. While far from conclusive, an argument can be made that the convergence of more basic cognitive processes and facilitator skills best explain the emergence of the factor structure in the sample.

Second Analysis

The specific question addressed by the second analysis was;

Can the latent variables derived from factors in the first analysis predict observer reported dysfunction in cognitive executive processes such as working memory, organization, planning, and self-monitoring?

In this analysis, the research study sought to ascertain if observed dysfunction in executive processes could be predicted by the latent EF factors in the sample. Although research has failed to find a direct relationship between scores on EF rating scales and formal assessment, it was thought that by separating groups into participants rated as having no discernible executive difficulties, and those who do, observer ratings and assessment would converge. Consistent with previous research, this convergence of formal assessment and observer ratings was not present in the current sample. Factors 2 and 3 revealed no ability to distinguish between groups in the full sample or among participants reporting as Caucasian. So inconsequential was the relationship between these factors and observed ratings, that they had a higher correlation to gender than to observed EF skills. When the Caucasian identified group was analyzed separately from other ethnic groups and the portion of the sample where ethnicity was not identified, Factor 1 did show a significant, albeit small, ability to predict group membership. However, further analysis indicated that, although multiple variables such as the D-KEFS Card Sort, NEPSY-II Animal Card Sort, and the WJIII COG NU Numbers Reversed showed significant correlations with group membership, only one variable, WJIII COG NU Concept Formation, emerged as a significant predictor of group membership. In both cases, the problem was primarily one of over-classification. Over 80% of the full sample rated as unimpaired by parents/guardians had EF assessment scores indicating they would experience dysfunction in executive skills. At first glance, it would appear that tasks severely over inflated estimates of difficulties; however, this does not take into consideration important factors. The first may have been in the use of parent ratings of EF. Observer ratings of EF are not absolute. In other words, rating scales are not measuring absolute abilities, but; rather, the perceived demonstration of a skill or trait in response to the demand of a specific environment. Thus, parent ratings of executive difficulties would be heavily dependent on the demands of the home environment, which may not tax EF skills to the same extent as the academic environment of school. It should be noted that the sample was a clinical sample meaning that the vast majority of children had one or more diagnoses that affected school and/or social performance. It seems somewhat unusual that 40% of the students in the sample were rated as having no deficits in key learning processes affecting self-monitored and self-regulated learning. The first explanation is that factors not present in the current research would better differentiate school difficulties amongst the clinical sample. Factors such as variance in verbal cognitive ability and "hot" executive processes (behavioral inhibition, and emotional regulation were excluded from the analysis) may better explain patterns in assessment and/or parent ratings. A second possible explanation is, as Bishop (2008) postulates, rating scales and assessment tasks do not measure the same constructs. While direct EF assessment measures the application of cognitive and EF skills as applied to a specific novel task, rating scales measure an opinion as to what extent an individual's demonstrated skill is able to meet the demands of a specific environment. In fact, the BRIEF exhibits moderate to high correlations with rating scales measuring emotional and behavioral difficulties. Even though rating scales are normed instruments, they are still survey measures and therefore do not assess innate abilities but rather, they express an opinion regarding the demonstration of those abilities. There is also the problem of objectivity. When completing a rating scale, a parent or teacher, in effect, is making a behavioral judgment about the child skills. It can be argued that a parent with a child experiencing cognitive and behavioral difficulty would therefore approach the task in a manner similar to a parent whose child did not experience these difficulties.

An analogy to the relationship between the dissimilarity between constructs assessed by formal measures and rating scales can be drawn with the assessment and rating of consumer products. One method of assessing the quality of products in a given category is to first establish what qualities or constructs are optimal in the product through consumer surveys. Using that data, a rating scale can be developed on which people can rate their satisfaction with a given product in that category (e.g. overall quality, durability, life span etc.). Another method would be to take multiple measurements within a sample of products to generate normative data to develop means and standard deviations of the product's quality in different conditions. This data can then be used to create test conditions to see how a given product compares to the normative sample. The first method approximates a rating scale while the second constitutes standardized, controlled assessment. There are drawbacks to both methods; however, the example clearly illustrates the two methods are not measuring the same constructs and are affected by significantly different factors. For instance, the consumer rating of a product would be highly affected by numerous factors including the consumer's expectations of performance, the degree to which the consumer provided care and maintenance, and any difficulties caused to the consumer because expectations of performance were not met. This analogy highlights some very salient issues with the determination of disabilities using rating scales as a primary measure of dysfunction.

Conclusions

The results of this research show that the structure of EF in a clinical population of school aged children approximates the factors that have been proposed in research; however, whether due to the clinical nature of the sample, or to other factors, certain theorized skills were harder to separate such as, working memory and concept formation/reasoning. In addition, the relationship between the skills tended to remain flat regardless of age, something that is counterintuitive to the development of EF in adolescence. In the final analysis, two strong factors emerged, each comprising an intersection of skill sets with the first being the skills of cognitive flexibility/concept formation and working memory necessary for drawing inferences and solving novel problems; and the second being a combined sustained/selective attention dimension that subsumed tasks of set shifting and inhibition. A third factor, possibly representing an aspect of self-monitoring and task strategy planning, did emerge but was mainly composed of the D-KEFS Tower task. Still, it was integral to the cohesiveness of the model showing the Tower task made an important contribution to the model. The way in which the factors emerged was best represented by the SNP/CHC Model (Miller, 2013) which postulates primary executive processes of reasoning, problem solving, inhibition and set shifting, supported by facilitators of lower and higher cognitive processes such as sustained and selective attention, working memory and processing speed.

In the final analysis, the latent factors were, overall, poor predictors of observer (parent) ratings of executive dysfunction. Factor 1 was a significant predictor of group membership for the reported ethnicity of Caucasian in the sample, but the effect was small and further analysis revealed only the WJIII COG NU Concept Formation task was significant in the model. Although the sample was composed of children with diagnoses that impaired social and scholastic achievement, only 57% were rated as having elevated levels of executive dysfunction, while 43% of parents stated that their child had average skills in monitoring, organization, planning, task initiation, and working memory. The assessment results indicated a higher level of dysfunction was present than was reported.

Limitations of the Study

Use of Archival Data

By its nature, the collection of archival data cannot be controlled or monitored by the researcher. The proposed study used archival data collected over a period of five to six years making time effects a possible threat to validity. In addition, collection of data over time required multiple graduate students from different cohorts as well as revisions of the database based on changes in theory and revisions of tests. To minimize the threat to inaccurate data, multiple safeguards were implemented with graduate students such as training and supervision in data entry and routine reviews of entered data by more experienced peers.

Although the individuals collecting the data were trained professionals under the supervision of the KIDS Inc. program, the test batteries and variables reported were chosen by the test administrators, limiting the number and types of variables used in the research. The use of archival data also limited the choice of available observer rating measures to the BRIEF, which was the only rating scale used with enough frequency for a sufficient sample to be drawn. As mentioned previously, the BRIEF's lack of demographic stratification in the normative sample has drawn criticism, and, despite subsequent research cited by the authors, presents a valid concern with the validity of this instrument.

Although the archival dataset presents with obvious disadvantages, it also affords some advantages that researcher collection of data would lack. For instance, the sample is representative of a larger geographic area, covering multiple states and geographic regions throughout the United States. In addition, the dataset includes a substantial cross section of the most commonly used assessment instruments in school-aged populations in a large clinical sample of children.

Data Imputation

Data imputation is an accepted and valid method of replacing missing data even in situations where upwards of 50% of data points are missing (Horton & Lipsitz, 2001). However, the fact that the statistical analyses were performed on data that was partly generated mathematically and not derived from actual participant scores must be noted. Because of the variability in test administration and reported scores among practitioners collecting the data, 45% of the data ultimately used in the study were imputed. Most variables fell between 40 and 50% missing data, with D-KEFS Trail Making having the least amount of data missing (28%) and WJIII COG NU Analysis Synthesis having the most (55%). By eliminating subtests with a large percentage of missing data and using advanced methods of data imputation, the integrity of the original data was maintained. However, less missing data might have resulted in a sample with different statistical parameters than the one that was imputed.

Clinical Population

While the use of a sample of children with mixed clinical diagnoses provides specific advantages to the proposed study, diagnoses reported by practitioners cannot be corroborated or validated. In addition, primary, secondary, and tertiary disabilities were reported, creating a complex interaction of factors that simply could not be addressed in the present study. However, this does not pose a significant limitation to the study. The focus of the research is determining agreement between assessment and observer ratings of executive dysfunction regardless of the underlying cause of that dysfunction. Thus, clinical diagnoses are less pertinent than the expected presence of an executive processing deficit. Descriptive statistics indicated consistency among results for all but a few of the subtests used despite gender, ethnicity or age. The one unaccounted for factor was the significant difference in scores between age groups on the BRIEF, which was not addressed by the study.

Reported and Utilized Scores

As mentioned previously, in archival data, there is no control over what scores are reported. Practitioners may report only higher order scores, omitting scores such as error rates that would be pertinent to the study. As mentioned previously, no assessment instrument can assess an executive process without also assessing the lower cognitive skills being regulated. The D-KEFS and NEPSY-II seek to correct this by providing baseline speed of processing conditions in some tasks. If basic processing speed is found to be impaired, then poor scores in the more complex conditions may not be reflective of deficits in executive processes. For instance, if an individual scores poorly on the naming portion of a Stroop task, it is likely a poor performance on the inhibition portion is due to basic processing speed and not due to a deficit in inhibitory processes. This may have significantly skewed results, especially on tasks measuring inhibition and attentional switching. Contrast scores were not readily available in the database, nor were all conditions commonly reported in every case, making it impossible to rule out deficits in lower cognitive skills as underlying reasons for poor task performance. This may explain, to some degree, why the D-KEFS and NEPSY –II tasks had lower factor loadings than the processing speed tasks of the WJIII COG NU. It is possible that the variance not shared between these tasks measured a latent EF construct while the factor itself simply measured cognitive efficiency of rote information processing.

Use of Latent Variables

By their nature, latent variables are combinations of shared variance between a number of individual variables (Field, 2009). Factor loadings of individual variables are, by nature, less than 1.0; in other words, only a portion of the weight of the individual variable variances is explained by the latent variable. In some cases, the individual variable may cross-load, meaning it may load substantially on more than one factor in the analysis. Thus, any statistical analysis performed with latent variables is subject to limitations in explaining the actual relationship between each original individual variable, and the dependent variables. The reason latent variables were necessary from a practical standpoint is because the statistical error introduced by using numerous individual variables in the LR would prohibit any valid conclusions. From a theoretical standpoint, the study was primarily concerned with the predictability of the underlying latent variables and not with the actual test scores. However, overall total communalities for the variables were relatively low meaning that the extracted factors did not account for as much variance in the dataset as originally anticipated. Thus, measurement error in the latent variables could have significantly affected the results.

Directions for Future Research

The results present both the complex nature of executive processes and the similarities in factor structure across populations. The emergent structure did indicate expected constructs and exhibited the difficulty in differentiating these constructs given a child and adolescent clinical sample. The research failed to find convergence between observer ratings of executive dysfunction and rating scales. This was consistent with previous research regardless of the use of different statistical procedures. Perhaps different results would have been obtained if teacher ratings were used to better reflect the demands of a school environment; however, as illustrated previously, this convergence might not be possible due to the dissimilarity of constructs measured through observation, and through formal assessment.

Another possibility for lack of convergence was that the use of the Metacognitive Index of the BRIEF, instead of individual scales such as working memory, initiation, etc., may have masked relationships between the study variables and more specific measures of dysfunction. For instance, although the first factor, Concept Formation/Working Memory, was not directly related to the BRIEF Metacognitive Index, a relationship may have existed with the more specific measure of the BRIEF Working Memory subscale. The logistics of which BRIEF subscales were best represented by different factors and the number of analyses that would have to be completed were beyond the scope of the current research; however, exploring such relationships would help to better define similarities or differences in assessed and observed constructs. The proliferation of other rating scales measuring executive constructs, as well as an expected revision of the BRIEF, provide ample opportunity for continued exploration of these relationships.

Future research should also continue to discern which combination of common assessment instruments reveal valid factor constructs in both non-clinical and clinical samples of children. It is important to research assessment tasks commonly used in practice to determine how assessed constructs differ between child and adolescent populations in both clinical and non-clinical populations. In addition, contrasts should continue to be made between the predictability of academic and social dysfunction using both observer ratings and formal assessment. Perhaps one or the other may prove to be a better predictor in certain situations or with certain populations. In the meantime, practitioners should be careful to address differences in rating scales and formal assessment using qualitative data to determine how these discrepancies can be explained.

Final Thoughts

The primary goals of the current study were to explore the factor structure of executive functions within a clinical sample and to see if emergent latent factors representing those functions could predict observed levels of impaired and normal executive competencies. The tasks used to examine EF within the sample were derived from test batteries commonly employed by school based and clinical practitioners; the WJIII COG NU, NEPSY-II, and D-KEFS. Tasks chosen were versions of tasks in various categories that showed empirical and theoretical validity in identifying executive function disorders.

Preliminary analyses showed that the clinical sample means and standard deviations were generally lower than those in the standardization sample; however, most mean variable scores still fell within the average range. These results are consistent with clinical studies reported in the NEPSY-II validity studies (Korkman et al., 2007b). Given these findings, practitioners should use caution in ruling out the diagnosis of a disability associated with executive dysfunction when tasks fall within normal limits. Rather, patterns of consistently lower scores and qualitative behavioral observations during task performance should also be considered in diagnosing a disability.

In the first analysis, similar tasks, for instance, tasks involving reasoning and categorical thinking skills generally exhibited moderate to strong correlations across batteries. However, in previous research, the factor structure between the D-KEFS and WJIII COG NU differed markedly (Dombrowski et al., 2013; Floyd et al., 2010), mostly due to the inclusion of verbal and visual spatial tasks. Reasoning tasks from the D-KEFS tended to cluster with tasks of comprehension and knowledge (*Gc*) from the WJIII COG NU. The factor structure found in the current research sought to minimize the effects of previously learning and therefore better approximated factors found in previous research. Practitioners may benefit from analyzing whether scores on verbally loaded tasks differ considerably from the tasks in this study comprising factors of EF. This may help discern whether a lack of previous knowledge or deficits in speech and language rather than innate deficits in executive processes underlie academic or behavioral difficulties.

Three factors emerged from the analysis representing categorical thinking/problem solving and working memory, selective attention and response inhibition/shifting, and planning/self-monitoring. Two of the three derived factors had multiple strong factor loadings that indicated they were valid constructs. The third factor had one main loading comprised of the D-KEFS Tower task and secondary loadings from tasks involving selective attention and working memory. Although this factor was weakly represented, the inclusion of the Tower task was necessary to preserve the emergent factor structure. The two primary factors were similar to previous research except for a lack of differentiation between related skills of working memory and reasoning, and between selective attention and response inhibition/set shifting. Overall, the results of the study support the importance of primary and secondary skills in assessing executive functioning. Broad problem solving and reasoning skills require facilitator skills such as working memory, focused attention, and effective information retrieval. Complex executive skills such as response inhibition and set shifting are heavily dependent on selective attention and speed of information processing. The results off the study seem to validate the approach of the D-KEFS and NEPSY-II, indicating that higher order skill deficits can only be ascertained by first assessing the efficacy of more basic cognitive processes. The SNP/CHC model highlights specific additional relationships between facilitator skills and higher order executive processes (e.g. working memory and concept formation). Clinicians are therefore urged to consider deficits in more basic cognitive skills such as short term/working memory and attentional processes in determining if they underlie or contribute to deficits in executive processes such as categorical thinking and inferential reasoning.

In the final analysis, the emergent EF factors were not predictive of observer ratings of executive function versus executive dysfunction as defined by significant or non-significant parent ratings on the BRIEF metacognitive index. Assessments tended to overestimate parent ratings of impairment. Factor 1 did predict observer ratings in the Caucasian group due largely to the effect of WJIII COG NU Concept Formation. Given the body of research failing to find a convergence of assessment and observer ratings, it is easy to "take sides" and favor one approach over the other. This author proposes another approach. Rather than looking for convergence, practitioners should understand that, much like different cognitive factors are assessed by different tasks, so do rating scales and tasks of EF measure different aspects of cognitive and behavioral functioning. Clinicians can use this difference to their advantage by specifically differentiating between observed/demonstrated competencies as assessed by rating scales and assessed/expected skills as determined by formal assessment. Using a systems approach, clinicians can then speculate as to why a skill might exist in assessment, and yet, not be demonstrated in one or more given settings. Likewise, it may also explain why the demands of a given environment may not tax an assessed deficit in function. If the ultimate purpose of assessment is to determine effective interventions and accommodations to facilitate academic and social success, such an approach would go far in determining appropriate and effective interventions for children with disabilities.

REFERENCES

Antzoulatos, E. G., & Miller, E. K. (2011). Differences between neural activity in prefrontal cortex and striatum during learning of novel abstract categories. *Neuron*, *71*, 243–249. http://doi.org/doi:10.1016/j.neuron.2011.05.040

Ardila, A. (1992). Luria's approach to neuropsychological assessment. *International Journal of Neuroscience*, *66*, 35–43. http://doi.org/10.3109/00207459208999787

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, *15*, 31–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14590565

Arnsten, A. F. T., & Li, B. M. (2005). Neurobiology of executive functions: Catecholamine influences on prefrontal cortical functions. *Biological Psychiatry*, *57*, 1377–1384.

Aron, A. R. (2008). Progress in executive-function research: From tasks to functions to regions to networks. *Current Directions in Psychological Science*, *17*, 124–129. http://doi.org/10.1111/j.1467-8721.2008.00561.x

Avirett, E. (2011). *Validity of executive functioning tasks across the WJ III COG, NEPSY, and D-KEFS in a clinical population of children: Applicability to three neurocognitive theories*. (Doctoral Dissertation) Texas Woman's University, Denton, TX. Retrieved from Proquest Dissertations and Theses database. (UMI No. 3464570)

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423. http://doi.org/10.1016/S1364-6613(00)01538-2

Baldo, J. V., Shimamura, A. P., Delis, D. C., Kramer, J., & Kaplan, E. (2001). Verbal and design fluency in patients with frontal lobe lesions. *Journal of the International Neuropsychological Society*, 586–596. http://doi.org/10.1017/S1355617701755063

Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*, 89–94. http://doi.org/10.1111/j.1467-8721.2009.01615.x

Banich, M. T., & Compton, R. J. (2011). *Cognitive neuroscience*. Belmont, CA: Wadsworth, Cengage Learning.

Bear, M. F. (2003). Bidirectional synaptic plasticity: From theory to reality. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*, 649–655. http://doi.org/10.1098/rstb.2002.1255

Beauchamp, M., Catroppa, C., Godfrey, C., Morse, S., Rosenfeld, J. V, & Anderson, V. (2011). Selective changes in executive functioning ten years after severe childhood traumatic brain injury. *Developmental Neuropsychology*, *36*, 578–595. http://doi.org/10.1080/87565641.2011.555572

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision-making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10731224

Benton, A. L. (2000). *Exploring the history of neuropsycology: Selected papers*. New York, NY: Oxford University Press. Retrieved from ebrary.com

Berretta, N., Nisticò, R., Bernardi, G., & Mercuri, N. B. (2008). Synaptic plasticity in the basal ganglia: A similar code for physiological and pathological conditions. *Progress in Neurobiology*, *84*, 343–362. http://doi.org/10.1016/j.pneurobio.2007.12.004

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*, 327–336. http://doi.org/10.1016/j.lindif.2011.01.007.Relations

Bishop, T. L. (2010). *Relationship between performance-based measures of executive function and the Behavior Rating Inventory of Executive Functioning (BRIEF), a parent rating measure*. (Doctoral Dissertation). Illinois Institute of Technology, Chicago. Retrieved from Proquest Dissertations and Theses (UMI No. 3435817)

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *The Behavioral and Brain Sciences*, *29*, 109–125. http://doi.org/10.1017/S0140525X06009034

Blair, C., Granger, D., Willoughby, M., Mills-Koonce, R., Cox, M., Greenberg, M. T., … Fortunato, C. K. (2011). Salivary cortisol mediates effects of poverty and parenting on executive functions in early childhood. *Child Development*, *82*, 1970–1984. http://doi.org/10.1111/j.1467-8624.2011.01643.x

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

Bouret, S., & Richmond, B. J. (2010). Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *The Journal of Neuroscience*, *30*, 8591–8601. http://doi.org/10.1523/JNEUROSCI.0049-10.2010

Broca, P. (2011). On the site of the faculty of articulated speech (1865). *Neuropsychology Review*, *21*, 230–235. http://doi.org/doi:10.1007/s11065-011-9173-y

Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of "hot" and "cool" executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, *24*, 337–349. http://doi.org/10.1016/j.ecresq.2009.06.001

Brooks, B. L., Sherman, E. M. S., & Strauss, E. (2009). NEPSY-II: A Developmental Neuropsychological Assessment, Second Edition. *Child Neuropsychology*, *16*, 80–101. http://doi.org/10.1080/09297040903146966

Brookshire, B., Levin, H., & Song, J. (2004). Components of executive function in typically developing and head-injured children. *Developmental Neuropsychology*, *25*, 61–83. http://doi.org/10.1080/87565641.2004.9651922

Brown, C. (2001). Descartes: Summary of some major points. Retrieved April 10, 2014, from trinity.edu/cbrown/modern/descartesMajorpoints.html

Brown, B., Hendrix, S.B., & Hedges, D. W. (2011). *Multivariate analysis for the behavioral and social sciences: A graphical analysis*. Hoboken, NJ: Wiley and Sons. Retrieved from http://site.ebrary.com

Brown, N. R., & Siegler, R. S. (1992). The role of availability in the estimation of national populations. *Memory & Cognition*, *20*, 406–412. http://doi.org/10.3758/BF03210924

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*, 273–293. http://doi.org/10.1207/S15326942DN1903_3

Burton, M. W., LoCasto, P. C., Krebs-Noble, D., & Gullapalli, R. P. (2005). A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *Neuroimage*, *26*, 647–661. http://doi.org/10.1016/j.neuroimage.2005.02.024

Cardoso-Leite, P., & Gorea, A. (2010). On the perceptual/motor dissociation: A review of concepts, theory, experimental paradigms and data interpretations. *Seeing and Perceiving*, *23*, 89–151. http://doi.org/10.1163/187847510X503588

Carlin, D. A. (2007). Gross morphology and architechtonics. In B. L. Miller & J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (2nd ed., pp. 92–106). New York, NY: The Guilford Press.

Carpenter, P. A., Just, M. A., & Reichle, E. D. (2000). Working memory and executive function: Evidence from neuroimaging. *Current Opinion in Neurobiology*, *10*, 195–199. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10753796

Chambers, C. D., Garavan, H., & Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neuroscience and Biobehavioral Reviews*, *33*, 631–646. http://doi.org/10.1016/j.neubiorev.2008.08.016

Chen, J. (2010). *Least Squares Method for factor analysis (Thesis)*. University of California. Retrieved from http://www.statistics.ucla.edu

Cheung, I., Shulha, H. P., Jiang, Y., Matevossian, A., Wang, J., & Weng, Z. (2010). Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proceedings of the National Academy of Sciences*, *107*(19), 2–7. http://doi.org/10.1073/pnas.1001702107

Collette, F., & Van der Linden, M. (2002). Brain imaging of the central executive component of working memory. *Neuroscience and Biobehavioral Reviews*, *26*, 105–125. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11856556

Coolidge, F. L., & Wynn, T. (2005). Working memory, its executive functions, and the emergence of modern thinking. *Cambridge Archaeological Journal*, *15*, 5–26. http://doi.org/10.1017/S0959774305000016

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 1–9. Retrieved from http://www.paronline.net

Coughlin, K. B. (2013). *An analysis of factor extraction strategies : A comparison of the relative strengths of principal axis, ordinary least squares, and maximum likelihood in research contexts that include both categorical and continuous variables*. (Doctoral Dissertation). University of South Florida. Retrieved from www.scholarcommons.usf.edu/etd/4459

Cramer, D. (2003). *Advanced quantitative data analysis*. Maidenhead, Berkshire, England: Open University Press: McGraw Hill Education.

Cummings, J. L., & Miller, B. L. (2007). Conceptual and clinical aspects of the frontal lobes. In B. L. Miller & J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (2nd ed., pp. 12–24). New York, NY: The Guilford Press. Retrieved from ebrary.com

Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York, New York: Avalon.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*, 2037–2078.

Davis, J. L., & Matthews, R. N. (2010). NEPSY-II Review: Korkman, M., Kirk, U., & Kemp, S. (2007). NEPSY--Second Edition (NEPSY-II). San Antonio, TX: Harcourt Assessment. *Journal of Psychoeducational Assessment*, *28*, 175–182. http://doi.org/10.1177/0734282909346716

De Pisapia, N., Slomski, J. A., & Braver, T. S. (2007). Functional specializations in lateral prefrontal cortex associated with the integration and segregation of information in working memory. *Cerebral Cortex*, *17*, 993–1006. http://doi.org/10.1093/cercor/bhl010

Delis, D., Kaplan, E., & Kramer, J. (2001a). *Delis Kaplan Executive Function System technical manual*. San Antonio, TX: The Psychological Corporation.

Delis, D., Kaplan, E., & Kramer, J. (2001b). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX.

DeLong, M., & Wichmann, T. (2010). Changing views of basal ganglia circuits and circuit disorders. *Clinical EEG and Neuroscience : Official Journal of the EEG and Clinical Neuroscience Society (ENCS)*, *41*, 61–67. http://doi.org/10.1177/155005941004100204

Denham, S., Warren-Khot, H. K., Bassett, H. H., Wyatt, T., & Perna, A. (2012). Factor structure of self-regulation in preschoolers: Testing models of a field-based assessment for predicting early school readiness. *Journal of Experimental Child Psychology*, *111*, 386–404. http://doi.org/10.1016/j.jecp.2011.10.002

Dickstein, S. G., Bannon, K., Castellanos, F. X., & Milham, M. P. (2006). The neural correlates of attention deficit hyperactivity disorder: An ALE meta-analysis. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*, 1051–1062. http://doi.org/10.1111/j.1469-7610.2006.01671.x

Dietert, R. R., Dietert, J. M., & Dewitt, J. C. (2011). Environmental risk factors for autism. *Emerging Health Threats Journal*, *4*, 1–12. http://doi.org/10.3402/ehtj.v4i0.7111

Distefano, C., & Mîndril, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, *14*(20), 1–11. Retrieved from http://pareonline.net/pdf/v14n20.pdf

Dombrowski, S. C., & Watkins, M. W. (2013). Exploratory and higher order factor analysis of the WJ-III full test battery: A school-aged analysis. *Psychological Assessment*, *25*, 442–455. http://doi.org/10.1037/a0031335

Donders, J., DenBraber, D., & Vos, L. (2010). Construct and criterion validity of the Behaviour Rating Inventory of Executive Function (BRIEF) in children referred for neuropsychological assessment after paediatric traumatic brain injury. *Journal of Neuropsychology*, *4*(Pt 2), 197–209. doi:10.1348/174866409X478970

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475–483. http://doi.org/10.1016/S0166-2236(00)01633-7

Ewing-Cobbs, L., Barnes, M. A., & Fletcher, J. M. (2003). Early brain injury in children: Development and reorganization of cognitive function. *Developmental Neuropsychology*, *24*, 669–704. http://doi.org/10.1080/87565641.2003.9651915

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17695343

Faw, B. (2003). Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition*, *12*, 83–139. http://doi.org/10.1016/S1053-8100(02)00030-2

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage Publishing.

Floyd, R. G., Bergeron, R., Hamilton, G., & Parra, G. R. (2010). How do executive funcitons fit with the Cattell-Horn-Carroll Model? Some evidence from a a joint factor analysis. *Psychology in the Schools*, *47*, 721–738. http://doi.org/10.1002/pits.20500

Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychological Science*, *17*, 700–707. http://doi.org/10.1111/j.1467-9280.2006.01769.x

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology*, *137*, 201–225. http://doi.org/10.1037/0096-3445.137.2.201

García-Madruga, J. A., Vila, J. O., Gómez-Veiga, I., Duque, G., & Elosúa, M. R. (2014). Executive processes, reading comprehension and academic achievement in 3th grade primary students. *Learning and Individual Differences*, *35*, 41–48. http://doi.org/10.1016/j.lindif.2014.07.013

Garcia-Molina, A. (2012). Phineas Gage and the enigma of the prefrontal cortex. *Neurologia*, *27*, 370–375. http://doi.org/10.1016/j.nrl.2010.07.015

Garlick, D., & Sejnowski, T. J. (2006). There is more to fluid reasoning then working memory and executive function. *Brain and Behavioral Sciences*, *29*, 134–135. http://doi.org/10.1017/S0140525X0630903X

Garrett, B. (2009). *Brain and behavior*. Los Angeles, CA: Sage.

Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders*, *5*, 143–154. http://doi.org/10.1177/108705470200500302

Gertsman, B. (2015). Unofficial website for basic biostatistics for public health practice. Retrieved August 12, 2014, from https://docs.google.com/document/d/1eDsev0OTjtqS1awmFc6LC36qbKjVIW8nT2P2Y4m6rtE/edit

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000a). [Test Review of the Behavior Rating Inventory of Executive Function (BRIEF), 2000]. *Child Neuropsychology*, *6*, 235–238. http://doi.org/10.1076/chin.6.3.235.3152

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

Gioia, G., Isquith, P. K., Retzlaff, P. D., & Espy, K. (2002). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychology*, *8*, 249–257. http://doi.org/10.1076/chin.8.4.249.13513

Goldberg, E. (2001). *The executive brain: Frontal lobes and the civilized mind*. New York, NY: Oxford Press.

Green, C. D. (2009). Darwinian theory, functionalism, and the first American psychological revolution. *The American Psychologist*, *64*, 75–83. http://doi.org/10.1037/a0013338

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. PNAS, 100, 253-258. doi:10.1073/pnas.0135058100. *Proceedings of the National Academy of Sciences*, *100*, 253–258. http://doi.org/10.1073/pnas.0135058100

Hackman, D. A., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*, 651–659. http://doi.org/10.1038/nrn2897

Hadjikhani, N. (2007). Mirror neuron sustem and autism. In P. C. Carlisle (Ed.), *Progress in autism research* (pp. 151–166). Hauppauge, New York: Nova Science Publishers.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. . (2010). *Multivariate data analysis*. Englewood Cliffs, CA: Prentice Hall.

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practioner's handbook*. New York, New York: The Guilford Press.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9*, 139–164. http://doi.org/10.1177/014662168500900204

Hebb, D., & Penfield, W. (1940). Human behavior after extensive bilateral removal from the frontal lobes. *Archives of Neuropsychiatry*, *44*, 421–428. http://doi.org/10.1001/archneurpsyc.1940.02280080181011

Henson, R. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*, 393–416. http://doi.org/10.1177/0013164405282485

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*, 26–32. http://doi.org/10.1016/j.tics.2003.11.003

Hochberg, Z. (2011). Developmental plasticity in child growth and maturation. *Frontiers in Endocrinology*, *2*, 1–6. http://doi.org/10.3389/fendo.2011.00041

Homack, S., Lee, D., & Riccio, C. A. (2005). [Review of the Delis-Kaplan Executive Function System (D-KEFS) by D. Delis, E. Kaplan & J. Kramer, 2003]. *Journal of Clinical and Experimental Neuropsychology*, *27*, 599–609. http://doi.org/10.1080/13803390490918444

Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, *28*, 617–644. http://doi.org/10.1207/s15326942dn2802_4

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*, 53–60.

Horton, N. J., & Lipsitz, S. R. (2001). Multiple imputation in practice : Comparison of software packages for regression models with missing variables. *The American Statistician*, *55*, 244–254. http://doi.org/10.1198/000313001317098266

Howse, R. B., Lange, G., Farran, D. C., & Boyles, C. D. (2003). Motivation and self-regulation as predictors of achievement in economically disadvantaged young children. *The Journal of Experimental Education*, *71*, 151–174. http://doi.org/10.1080/00220970309602061

Hurley, A., & Scandura, T. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior*, *18*, 667–683. Retrieved from https://wweb.uta.edu/management/Dr.Casper/Fall10/BSAD6314/Coursematerial/Hurley et al 1997 - Exploratory and Confirmatory Factor Anal.pdf

Illig, D. C. (1998). *Birth to Kindergarten: The importance of the early years*. *California Research Bureau* (Vol. 95814). Retrieved from www.library.ca.gov

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

Kailasanath, V., & Fu, S. (2010). The Hopes brain tutorial. Retrieved from group/hopes/cgi-bin/wordpress/2010/07/the-hopes-brain-tutorial-flash-version/

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children, Second Edition*. Circle Pines, MN: American Guidance Service.

Kofler, M. J., Rapport, M. D., Bolden, J., Sarver, D. E., Raiker, J. S., & Alderson, R. M. (2011). Working memory deficits and social problems in children with ADHD. *Journal of Abnormal Child Psychology*, *39*, 805–817. http://doi.org/10.1007/s10802-011-9492-8

Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances: Is a Heywood Case a symptom of misspecification? *Sociological Methods & Research*, *41*, 124–167. http://doi.org/10.1177/0049124112442138

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

Korkman, M., Kirk, U., & Kemp, S. L. (2007b). *NEPSY II: A Developmental Neuropsychological Assessment. Clinical and interpretive manual*. San Antonio, TX: Psychological Corporation.

Kortte, K. B., Horner, M. D., Carolina, S., & Windham, W. K. (2002). The Trail Making Test , Part B : Cognitive flexibility or ability to maintain set ? *Applied Neuropsychology*, *9*, 106–109. http://doi.org/10.1207/S15324826AN0902_5

Kowalska, D. M., Bachevalier, J., & Mishkin, M. (1991). The role of the inferior prefrontal convexity in performance of delayed nonmatching-to-sample. *Neurpsychologica*, *29*, 583–600. http://doi.org/10.1016/0028-3932(91)90012-W

Languis, M., & Miller, D. (1992). Luria's Theory of brain functioning: A model for research in cognitive psychology. *Educational Psychologist*, *27*, 493–511. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15326985ep2704_6#preview

Latzman, R. D., Elkovitch, N., Young, J., & Clark, L. A. (2010). The contribution of executive functioning to academic achievement among male adolescents. *Journal of Clinical and Experimental Neuropsychology*, *32*, 455–462. http://doi.org/10.1080/13803390903164363

Latzman, R. D., & Markon, K. E. (2010). The factor structure and age-related factorial invariance of the Delis-Kaplan Executive Function System (D-KEFS). *Assessment*, *17*, 172–184. http://doi.org/10.1177/1073191109356254

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

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

Leonard-Zabel, A., & Feifer, S. G. (2009). Frontal lobe dysfunction, psychopathology and violence. In S. G. Feifer & G. Rattan (Eds.), *Emotional disorders: A neuropsychological, pharmacological and educational approach* (pp. 107–122). Middleton, MD: School Neuropsych Press, LLC.

Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology*, *17*, 281–297. http://doi.org/10.1080/00207598208247445

Linder, T. (2008). *Transdisciplinary play-based intervention 2: Guidelines for developing a meaningful curriculum for young children*. Towson, MD: Paul H. Brooks Publishing.

Luria, A. R. (1973). *The working brain*. New York, NY: Basic Books.

Luria, A. R. (2010). L . S . Vygotsky and the problem of functional localization. *Journal of Russian and East European Psychology*, *40*, 17–25. http://doi.org/10.2753/RPO1061-0405050353

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, *36*, 611–637. http://doi.org/10.1207/S15327906MBR3604_06

Mahone, E. M., Cirino, P. T., Cutting, L. E., Cerrone, P. M., Hagelthorn, K. M., Hiemenz, J. R., … Denckla, M. B. (2002). Validity of the Behavior Rating Inventory of Executive Function (BRIEF) in children with ADHD and/or Tourette syndrome. *Archives of Clinical Neuropsychology*, *17*, 643–662. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14591848

Mareschal, D., & Schultz, T. R. (1999). Development of children's seriation: A connectionist approach. *Connection Science*, *11*, 149–186.

Mattison, R. E., & Mayes, S. D. (2012). Relationships between learning disability, executive function, and psychopathology in children with ADHD. *Journal of Attention Disorders*, *16*, 138–146. http://doi.org/10.1177/1087054710380188

McCabe, D. P., Roediger, H. L., Mcdaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*, 222–243. http://doi.org/10.1037/a0017619

McCloskey, G., Hewitt, J., Henzel, J. N., & Eusebio, E. (2009). Executive functions and emotional disturbance. In S. Feifer & G. Rattan (Eds.), *Emotional disorders: A neuropsychological, psychopharmalogical, and educational perspective* (pp. 65–106). Middleton, MD: School Neuropsych Press, LLC.

McCloskey, G., Perkins, L. A., & Van Divner, B. (2006). *Assessment and intervention for executive function difficulties*. New York, NY: Tayor and Francis Group, LLC.

McCrimmon, A. W., Schwean, V. L., Saklofske, D. H., Montgomery, J. M., & Brady, D. I. (2012). Executive functions in Asperger's syndrome: An empirical investigation of verbal and nonverbal skills. *Research in Autism Spectrum Disorders*, *6*, 224–233. http://doi.org/10.1016/j.rasd.2011.05.003

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10. http://doi.org/10.1016/j.intell.2008.08.004

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Technical manual, Woodcock-Johnson III Normative update*. Rolling Meadows, IL: Riverside Publishing.

Meaney, M. J. (2010). Epigenetics and the biological definition of gene x environment interactions. *Child Development*, *81*, 41–79. http://doi.org/10.1111/j.1467-8624.2009.01381.x

Meltzer, L., & Krishman, K. (2007). Executive function difficulties and learning disabilities: Understandings and misunderstandings. In L. Meltzer (Ed.), *Executive function in education* (pp. 77–105). New York, NY: The Guilford Press.

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak Publishing.

Mesulam, M. M. (2002). The human frontal lobes: Transcending the default mode through contingent encoding. In D. Stuss & R. Knight (Eds.), *The human frontal lobes* (pp. 8–30). New York, NY: Oxford Univeiisty Press.

Milberg, W. P., Hebben, N., & Kaplan, E. (2009). The Boston process approach to neuropsychological assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric and neuromedical disorders* (3rd ed., pp. 42–65). New York, NY: Oxford Univeiisty Press.

Miller, D. (2013). *Essentials of school neuropsychological assessment* (3rd ed.). Hoboken, NJ: Wiley and Sons.

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

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment (2nd ed.)*. New York, NY: Oxford University Press.

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

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*, 49–100. http://doi.org/10.1006/cogp.1999.0734

Müller, N. G., & Knight, R. T. (2006). The functional neuroanatomy of working memory: Contributions of human brain lesion studies. *Neuroscience*, *139*, 51–58. http://doi.org/10.1016/j.neuroscience.2005.09.018

Munakata, Y., & McClelland, J. L. (2003). Connectionist models of development. *Developmental Science*, *6*, 413–429. http://doi.org/10.1111/1467-7687.00296

Munataka, Y., Mclelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review*, *104*, 686–713. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9337629

Naglieri, J. A., & Rojahn, J. (2004). Construct validity of the PASS Theory and CAS: Correlations with achievement. *Journal of Educational Psychology*, *96*, 174–181. http://doi.org/10.1037/0022-0663.96.1.174

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self regulation* (pp. 1–18). New York, NY: Plenum.

Osborne, J. W. (2013). Dealing with missing or incomplete data: Debunking the myth of perfect data. In *Best practices in data cleaning* (pp. 105–132). Thousand Oaks, CA: Sage Publishing.

Packwood, S., Hodgetts, H. M., & Tremblay, S. (2011). A multiperspective approach to the conceptualization of executive functions. *Journal of Clinical and Experimental Neuropsychology*, *33*, 456–470. http://doi.org/10.1080/13803395.2010.533157

Pallier, G., Roberts, R. D., & Stankov, L. (2000). Biological versus psychometric intelligence: Halstead's (1947) distinction revisited. *Archives of Clinical Neuropsychology*, *15*, 205–226. http://doi.org/10.1093/arclin/15.3.205

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews*, *2*, 417–424. http://doi.org/10.1038/35077500

Phillips, L. H., Wynn, V., Gilhooly, K. J., Della Sala, S., & Logie, R. H. (1999). The role of memory in the Tower of London task. *Memory & Cognition*, *7*, 209–231. http://doi.org/10.1080/741944066

Plucker, J. (2014). Wilhelm Wundt. Retrieved from Http://www.indiana.edu/~intell/wundt.shtml

Powell, K. B., & Voeller, K. K. S. (2004). Prefrontal executive function syndromes in children. *Journal of Child Neurology*, *19*, 785–797. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15559894

Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model selection perspective. *Multivariate Behavioral Research*, *48*, 28–56. http://doi.org/10.1080/00273171.2012.710386

Purdy, M. H. (2005). Executive Functions: Theory, assessment , and treatment. In M. Kimbarrow (Ed.), *Cognitive communication disorders*. San Diego, CA: Plural Publishing.

Purisch, A. D. (2001). Misconceptions about the Luria-Nebraska Neuropsychological Battery. *Neurorehabilitation*, *16*, 275–280. Retrieved from http://iospress.metapress.com

Rathbone, R., Counsell, S. J., Kapellou, O., Dyet, L., Kennea, N., Hajnal, J., … Edwards, A. D. (2011). Perinatal cortical growth and childhood neurocognitive abilities. *Neurology*, *77*, 1510–1515. http://doi.org/10.1212/WNL.0b013e318233b215

Rattan, G. (2009). Psychopharmacology, depression and alternative treatment modalities. In S. G. Feifer & G. Rattan (Eds.), *Emotional disorders: A neuropsychological, psychopharmacological and educational perspective* (pp. 147–199). Middleton, MA: School Neuropsych Press, LLC.

Reitan, R. M. (1994). Ward Halstead's contributions to neuropsychology and the Halstead-Reitan Neuropsychological Test Battery. *Journal of Clinical Psychology*, *50*, 47–70. Retrieved from http://psycnet.apa.org/psycinfo/1994-24013-001

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery* (2nd Editio). Tucson, AZ: Neuropsychology Press.

Repovs, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*, 5–21. http://doi.org/10.1016/j.neuroscience.2005.12.061

Reynolds, C. R., Castillo, C. M., & Horton, A. R. (2007). Neuropsychology and intelligence: An overview. In H. A. MacNeill & D. Wedding (Eds.), *Neuropsychology Handbook* (pp. 69–86). New York, New York: Springer.

Roth, R. M., Erdodi, L. A., McCulloch, L. J., & Isquith, P. K. (2015). Much ado about norming: The Behavior Rating Inventory of Executive Function. *Child Neuropsychology*, *21*, 225–33. http://doi.org/10.1080/09297049.2014.897318

Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., … Coffey, C. E. (2002). Executive function: A review of its promise and challenges. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *14*, 377–405. http://doi.org/doi:10.1176/appi.neuropsych.14.4.377

Ruh, N., Cooper, R. P., & Mareschal, D. (2005). The time course of routine action. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *27th Annual Meeting of the Cognitive Science Society, July 21-23, 2005*. Stressa, Italy. Retrieved from http://eprints.bbk.ac.uk/614/

Sabbatini, R. M. . (2003). Brain localization in the 19th century. *Brain and Mind*. Retrieved from http://www.cerebromente.org.br/n01/frenolog/frenloc.htm

Schmitter, A. (2014). 17th and 18th century theories of emotions. In E. M. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring). Retrieved from http://plato.stanford.edu/archives/spr2014/entries/emotions-17th18th/

Schneider, W. J., & McGrew, K. (2012). The Catell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). New York: Guilford Press.

Schrank, F.A., Miller, D.C., Wendling, B.J., & Woodcock, R. W. (2010). *Essentials of WJIII cognitive abilities assessment* (2nd edition). Hoboken, New Jersey: Wiley and Sons.

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

Senn, T., Espy, K., & Kaufmann, P. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology*, *26*, 445–464. http://doi.org/10.1207/s15326942dn2601_5

Spencer-Smith, M., & Anderson, V. (2009). Healthy and abnormal development of the prefrontal cortex. *Brain*, *12*, 279–297. http://doi.org/10.1080/17518420903090701

Stringer, A. Y., & Cooley, E. L. (2002). Neuropsychology: A 20th century science. In A. Y. Stringer, E. L. Cooley, & A. Christensen (Eds.), *Pathways to prominence in neuropsychology: 20th century reflections* (pp. 1–26). New York, NY: Psychology Press. Retrieved from https://books.google.com

Stuss, D. T., & Knight, R. T. (Eds.). (2002). *Principles of frontal lobe function*. Cary, NC: Oxford University Press. Retrieved from http://www.ebrary.com

Stuss, D. T., & Levine, B. (2002). Adult clinical neuropsychology: Lessons from studies of the frontal lobes. *Annual Review of Psychology*, *53*, 401–433. Retrieved from http://www.arjournals.annualreviews.org

Taherdoost, H., Sahibuddin, S., & Jalaliyoon, N. (2004). Exploratory factor analysis: Concepts and theory. *Advances in Applied and Pure Mathematics*, 375–382.

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

Thimble, M. H. (1990). Psychopathology of frontal lobe syndromes. *Seminars in Neurology*, *10*(3), 1–6. Retrieved from http://www.ect.org/

Toll, S. W. M., Van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities*, *44*, 521–532. http://doi.org/10.1177/0022219410387302

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*, 131–143. http://doi.org/10.1111/jcpp.12001

Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *The British Journal of Developmental Psychology*, *31*, 70–87. http://doi.org/10.1111/j.2044-835X.2012.02079.x

Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Ebstein, R. P. (2011). Methylation matters in child development: Toward developmental behavioral epigenetics. *Child Development Perspectives*, *5*, 305–310. http://doi.org/10.1111/j.1750-8606.2011.00202.x

Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, *165*, 710–718. http://doi.org/10.1093/aje/kwk052

Watkins, M. W. (2006). Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods*, *5*, 344–346. Retrieved from http://www.public.asu.edu/

Weiner, I. B., Schinka, J. A., & Velicer, W. F. . (2012). *Handbook of psychology: Research methods in psychology* (2nd ed.). Somerset, NJ, USA: John Wiley & Sons. Retrieved from ProQuest ebrary. Web. 11 March 2015.

Wenzel, C. H. (2006). The sensus communis and its subjective aspects. From Aristotle and Cicero via Aquinas to Kant. In *Conference of the Taiwan Philosophical Association* (pp. 1–7).

White, R. F., & Rose, F. E. (1997). The Boston Process Approach: A brief history and current practice. In G. Goldstein & T. Incagnoli (Eds.), *Contemporary approaches to neuropsychological assessment* (pp. 171–211). New York, NY: Plenum Press.

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*, 418–431. http://doi.org/10.1037/a0025779

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *The Woodcock Johnson Tests of Cognitive Ability, 3rd Edition Normative Update (WJ-III, NU)*. *2001*. Itasca, IL: Riverside Publishing.

Wundt, W. (1910). *Principles of physiological psychology, Volume 1. (E. Tichener, Trans.)* (2nd ed.). London: Swan Sonnenschein.

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, *9*, 79–94. Retrieved from http://www.tqmp.org/RegularArticles.html

Zhang, Y., Gan, Y., Cham, H., Wang, M.-T., Willett, J. B., Eccles, J. S., … Appleton, J. J. (2012). School Engagement Trajectories and Their Differential Predictive Relations to Dropout. *Journal of Adolescence*, *74*, 274–283. http://doi.org/10.1002/pits

Zook, N. A., Davalos, D. B., Delosh, E. L., & David, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition*, *56*, 286–292. http://doi.org/http://dx.doi.org/10.1016/j.bandc.2004.07.003