8-9-2014

PSYCHOMETRIC EVALUATION OF AN EXECUTIVE FUNCTION BATTERY FOR PEDIATRIC SICKLE CELL DISEASE

Melita T. Stancil
University of South Carolina - Columbia

PSYCHOMETRIC EVALUATION OF AN EXECUTIVE FUNCTION BATTERY FOR PEDIATRIC SICKLE CELL DISEASE

by

Melita T. Stancil

Bachelor of Arts
Norfolk State University, 2005

Master of Arts
University of South Carolina, 2011

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Clinical-Community Psychology

College of Arts and Sciences

University of South Carolina

2014

Accepted by:

Jeffrey C. Schatz, Major Professor

Bret Kloos, Committee Member

Rhonda Jeffries, Committee Member

Jennifer C. Vendemia, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

DEDICATION

To my son, Jimmy M. Stancil II.

Words cannot express the gratitude that I will always feel for your sacrifice of my time during the completion of this degree.

ACKNOWLEDGEMENTS

I am appreciative for the support of the many individuals who made this dissertation possible. To the children and families who participated in this investigation, the medical team at Palmetto Health Richland, the community and school directors and their staff, the Neuropsychology and Human Development Laboratory research assistants, Caitlin Oliver and Tal Katz. I will always appreciate the willingness and dedication that was necessary to ensure that this study was successful. Thank you for your participation in this project.

I am also thankful for the guidance and mentorship of Dr. Jeffrey C. Schatz, my major professor and dissertation advisor. Matriculation through the program and achievement of this work would not have been possible without his continued encouragement and knowledge. The research and clinical training that I have received under his supervision have been invaluable. Thank you for sharing your knowledge and constructive feedback towards my professional development.

Furthermore, I would like to recognize my dissertation committee members, Drs. Bret Kloos, Jennifer C. Vendemia, and Rhonda Jeffries. Their understanding, collaboration, and direction were instrumental in guiding the cultivation of this manuscript. Thank you all for your flexibility and tutelage during this process.

ABSTRACT

Sickle Cell Disease (SCD) is a genetic disorder which affects hemoglobin and is associated with high rates of neurologic and neurocognitive deficits. Recent studies have indicated executive functioning (EF) as a common area of impairment for children diagnosed with SCD; however, there is no consensus about which measures of executive function are best to use in clinical practice or research. The purpose of the present research is to assess the properties of a new executive function measure, the "EXAMINER" to determine its utility with the SCD population. Thirty-two children with SCD and 86 demographically-matched controls completed established cognitive measures known to be sensitive to SCD-related neurocognitive deficits as well as EXAMINER tasks of executive attention, set shifting, working memory, inhibition, planning, and fluency. A statistical analysis compared performance on the attention, set shifting, inhibition, planning, and fluency measures in the EXAMINER relative to established measures (i.e., the verbal comprehension, processing speed, short term memory measures in the Woodcock Johnson Tests of Cognitive Abilities III) to compare reliability and validity. It was expected that the EXAMINER would show comparable reliability and validity to the established measures using traditional definitions of these constructs. In addition, examination of cultural validity was examined due to the high preponderance of SCD among the African-American population. Statistical analysis indicated support for internal consistency, convergent validity, cultural validity, and

sensitivity of the EXAMINER, relative to traditional cognitive measures. Taken together, these results suggest that the EXAMINER is a useful neuropsychological test for the pediatric SCD population. Future research should focus on using more sophisticated scale development approaches, including greater ethnic and economic diversity with norming groups, and incorporating children in the statistical models to further establish strong psychometric properties.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The identification of appropriate executive function (EF) measures for pediatric sickle cell disease (SCD) is pivotal in understanding disease related impacts on cognitive functioning. SCD is an autosomal recessive genetic blood disorder commonly found in individuals with African, Mediterranean, Indian, and Middle Eastern heritage. Approximately 1 in 11 African Americans in the United States (US) carry the sickle cell trait and it is estimated that 1 in 500 African-Americans have SCD in the United States (Charache, Lubin, & Reid, 1992).

SCD is a multi-systemic disease that ultimately impacts the biological integrity of organs, especially the brain. Decrements associated with SCD brain effects have been found in a wide range of cognitive areas that can be categorized as EF. EF is a broad construct that incorporates a collection of interconnected cognitive processes responsible for focused, goal oriented behavior (Gioia, Isquith, & Guy, 2001). Even when children with SCD appear normal by magnetic resonance imaging (MRI) scan and on traditional measures of cognition, such as Intelligence Quotient (IQ), cognitive functioning has been found to be impaired (Steen, Fineberg-Buchner, Hankins, Weiss, Prifitera, & Mulhern, 2005; Schatz, Finke, Kellett, & Kramer, 2002; Wang et al., 2001). However, pediatric SCD treatment studies have consistently utilized neurologic imaging technologies and global measures of IQ to understand cognition, as opposed to modality specific

neuropsychological assessment (Schatz, Finke, Kellett, & Kramer, 2002).

Methodological issues associated with the measurement of cognitive functioning are important several reasons, especially in this population. First, multi-center SCD randomized clinical treatment trials have been halted by The National Heart Lung and Blood (NHLB) for safety concerns associated with aparticipants completing neuroimaging exams (Lee, Piomelli, Granger, Miller, Harkness, Brambilla, Adams, 2006). Second, neuroimaging exams may only capture structural lesions and underestimate the full extent of neurocognitive deficits (Schatz, Brown, Pascual, Hsu, DeBaun, 2001; Schatz, Finke, Kellett, & Kramer, 2002). In addition, the brain health of children with SCD may be improved if measurement of the narrow abilities associated with cognition were incorporated in pediatric SCD treatment studies. Pediatric SCD treatment studies have excluded specific measures of cognitive ability performance from their outcomes, in part due to inconsistent measurement practices in the research literature. Deficits in children with SCD have been shown to impact general intellectual ability, attention, language, visual-spatial abilities, short and long term memory, and academic outcomes (Steen, Fineberg-Buchner, Hankins, Weiss, Prifitera, & Mulhern, 2005; Schatz, Finke, & Roberts, 2004; Schatz, Finke, Kellett, & Kramer, 2002). While this distinct cognitive profile for pediatric SCD is apparent, there is a lack of standard psychometric tools which evaluate these cognitive abilities across this body of literature. Furthermore, within this literature, there is a wide assortment of neuropsychological

measures which are used to investigate cognition partly due to the lack of systematic evaluation of the measures themselves for use with the SCD population.

The development of a neuropsychological assessment battery focused on executive functions that will meet or exceed the current established by cognitive batteries, such as the Wechsler Intelligence Scales for Children (WISC) and the Woodcock Johnson-III Tests of Cognitive Abilities (WJ-III), will help the field develop more accurate measurement of disease impacts for children with pediatric SCD. The issue of having well validated tools to measure executive functions for clinical research is common across multiple neurologic conditions. The Executive Abilities: Methods and Instruments for Neurobehavioral Evaluation and Research (EXAMINER) was constructed to address this limitation. The overall goal for the development of the EXAMINER was to reliably and validly assess domains of EF for clinical investigations and clinical trials that were adaptable across a wide range of ages and disorders. The purpose of the present investigation was to validate the psychometric integrity of the EXAMINER for use in children with SCD.

Organization of Review

This study sought to integrate SCD research, SCD treatment, and EF literature to propose utilization of a standard psychometric battery which will impact both science as well as practice. The background information to set the stage for this study was conducted in three separate sections. The first section provided background information on SCD and SCD associated neurologic complications. Pediatric SCD studies examining cognitive functioning in relation to neurologic disease and treatment were discussed. The

concept of EF, construct conceptualization, and measurement in the context of clinical outcomes and clinical decision making was reviewed in the second section.

The third section, of this document, assessed the current state of the science for cognitive tools in pediatric SCD literature and identified the minimal standards that EF measures should meet or exceed in order to become the new state of the art for clinical outcome studies in SCD. In this section, an overall description of the measures that are currently in use with this population is presented. Additionally, the psychometrics of specific tests utilized with pediatric SCD are reviewed. Strengths and weaknesses associated with the cognitive measures, used to-date, with the SCD population are also described. Finally, general comments are offered about the current state of EF measures within these existing studies.

Section I: Sickle Cell Disease

Nearly 50,000 Americans in the United States (US) have SCD and the majority of these individuals are of African, Mediterranean, Indian, and Middle Eastern descent (Edwards et al., 2005; McCrae & Lumley, 1998). It has been estimated that 1 out of 500 African Americans have a genetic variation of sickle cell disease (Charache, Lubin, & Reid, 1992). SCD refers to a category of inherited blood disorders (McCrae and Lumley, 1998). SCD occurs when an individual inherits an S-type hemoglobin (sickle type hemoglobin) gene along with another hemoglobin gene that does not code for A-type hemoblobin (the most common form of hemoglobin). Hebbel, Mohandas, Embury, & Hebbel (1994) describe four major genotypes of SCD: hemoglobin SS (HbSS), hemoglobin SC, (HbSC), HbS-beta-zero-thalassemia ($HBS\beta^0$), and HbS-beta-plus thalassemia ($HBS\beta^+$). Sickle Cell Anemia, hemoglobin SS, is the most severe form of the

disease and affects 65% to 75% of people diagnosed with SCD in the US (Platt, Orkin, Dover, Beardsley, Miller, Nathan, 1984).

Individuals with SCD produce abnormal hemoglobin, which is influenced by the genetic genotype of the disease. When oxygen saturation in the blood drops below a critical threshold the red blood cells (RBCs) outer membrane change from typical disk-like, “donut” shape to a curved, elongated, “banana” or “sickle” shape. Rigid “sickled” cells often lead to common complications associated with the disease, due to these cells having reduced lifespan and poorer oxygen carrying capacity (Serjeant, 1997). Chronic symptoms include unpredictable pain from vessel occlusion, anemia, increased fatigue, splenic sequestration, lung functioning impairment, priapism, enuresis, growth delays, increased risk of infection, tissue infarction, cerebral vascular disease (CVD), and acute chest syndrome (Sickle Cell Disease Guideline Panel, 1993).

Chronic symptoms related to SCD have been found to directly and indirectly impact cognitive ability through brain function. In particular, CVD and neurologic dysfunctions related to blood vessel function have been found to increase the probability of neurologic complication (Hillery & Panepinto, 2004). This damage is thought to occur because “sickled” cells reduce oxygen delivery to the brain and also damage small and large vessels in the brain (Huttenlocher, Mohr, Johns, et al. 1994). Silent stroke can be evidenced on clinical MRI exams in areas such as the arteriole-capillary-venule beds, such as found in white matter regions at the distal parts of the major cerebral arteries (Adams, McKie, Hsu, Files, Vichinsky, et al., 1998). Overt stroke, which is evidenced with clinical MRI exams and an observable neurologic event, can be present from the internal carotids to the circle of Willis (DeBaun, Derdeyn, McKinstry, 2006). In children

with SCD and overt stroke, decreased verbal intelligence quotient (VIQ), full-scale intelligence quotient (FIQ), and narrower abilities such as freedom from distractibility, attention, and other location-specific functions (e.g., motor skills contralateral to the injury) have been demonstrated (Craft, Schatz, Glauser, Lee, DeBaun, 1993; Steen, Emudianughe, Hankins, Wynn, et al., 2003). Less severe neurocognitive deficits, in similar areas of cognitive ability as overt stroke, have been associated with silent stroke in children with SCD (Armstrong, Thompson, Wang, Zimmerman, Pegelow, Miller, et al., 1996; Bernaudin, Verlhac, Freard, Roudot-Thoraval, Benkerrou, et al., 2000). Furthermore, deficits associated with FIQ, general cognitive ability, and attention-related skills (e.g., short term memory and processing speed) have been shown for those with SCD without identifiable brain injury on clinical MRI (Schatz et al., 2004). These deficits have been demonstrated to not be related to general disease severity (e.g., hospitalization rates, pain severity, anemia severity) and are above and beyond social risk factors for poorer cognitive test scores (for review, see Schatz et al., 2002). In terms of functional behavior, the broad and narrow cognitive processes affected by SCD are important as inter-related processes which are responsible for purposeful, goal-directed behavior. This is one of the definitions of EF (Gioia, Isquith, Guy, and Kenworthy, 2000).

SCD neurologic injury has been a focal point of treatment studies; however assessing EF from a psychological perspective is uncommon in pediatric SCD treatment studies. Many of these studies have not included cognitive measures due to contextual and methodological considerations in the research associated with pediatric SCD and cognition. Many pediatric hematology centers have limited access to psychology resources. The use of cognitive tools has been inconsistent across clinical studies.

Methodologically, in contrast to tools such as clinical/structural MRI exams for identifying cerebral infarcts, there is not a gold standard assessment battery recommended to measure cognitive deficits in this population. This limitation reduces the use of cognitive abilities as a primary measure of treatment outcomes. The absence of established, standard neuropsychological tests may also influence the results obtained for children with SCD. The cognitive performance of children with SCD has often been measured with tools idiosyncratic to each study, and many of which have varying psychometric quality. Additionally, research studies which evaluate cognition in pediatric SCD generally have small samples sizes, thus restraining the generalizability of findings (Nikhar et al., 2011). Furthermore, pediatric SCD research methodology is varied in nature which creates challenges in comparing findings across studies (e.g., different control groups, different approaches to sampling, etc.). Although there are limitations associated with pediatric SCD neuropsychological research, neuropsychological assessment is generally viewed as valid and reliable with validity measures equaling or exceeding those of medical tests, including neuroimaging (Matarazzo, 1990; Meyer, Finn, Eyde, et al., 2001). Therefore, the inclusion of cognitive measures is important for improved understanding of disease related mechanisms and treatments.

Overview of neurocognitive relationships in SCD

An examination of the existing literature on the cognitive effects of pediatric SCD is essential for the understanding how to best select neuropsychological measures to include in treatment studies. Stroke, silent infarct, and general disease status differentially affect the severity of cognitive deficits. Additionally, the cognitive performance of children has been shown to be impacted by location, size, and volume of neurologic

tissue damaged by SCD morbidity (Watkins, Hewes, Connelly, Kendall, Kingsley, Evans, Gadian, Vargha-Khadem, Kirkham, 1998). Compared to healthy children, stroke risk has been shown to increase by 221-fold in pediatric SCD (King, Herron, McKinstry, Bacak, Armstrong, White, and DeBaun, 2006). Overall, stroke has been shown to negatively impact language ability, attention, working memory, processing speed, and executive functioning on various neuropsychological tests (Steen et al., 2005; Schatz, Finke, & Roberts, 2004; Schatz, Finke, Kellett, & Kramer, 2002; Noll et al., 2001; Schatz et al., 2001). In addition, the SCD-related risk of silent cerebral infarct is 410-fold higher versus peers without SCD (King, Herron, McKinstry, Bacak, Armstrong, White, and DeBaun, 2006). Children diagnosed with SCD and silent cerebral infarcts generally perform poorer in FIQ, VIQ, performance intelligence quotient (PIQ), language, memory, academic achievement, attention, processing speed and motor coordination than children with SCD and no MRI abnormality (Bernaudin et al., 2000; Schatz et al. 2001; Schatz & Buzan, 2006; Steen et al. 2005). Furthermore, cognitive decrements have also been demonstrated in children with SCD who have no history of brain abnormalities evidenced on MRI (Schatz et al., 2002; Steen 2003; 2005). Overall, a wide range of psychological domains such as general intellectual ability, attention, language, visual-spatial abilities, short and long term memory, executive functioning, and academic outcomes have been shown to be reduced in children diagnosed with SCD when compared to healthy peers (Steen et al., 2005; Schatz, Finke, & Roberts, 2004; Schatz, Finke, Kellett, & Kramer, 2002).

SCD Treatment Studies

Most pediatric SCD treatment studies provide a limited view of cognitive effects associated with neurologic change from treatment therapies. Treatment studies traditionally seek to alleviate complications associated with pediatric SCD by relieving pain and managing overall disease related complications (Puffer, Schatz, Roberts, 2007). Additionally, the primary focus of this body of literature is the prevention of secondary neurologic injury after an initial brain insult has occurred, due to SCD complications (Adams, Ohene-Frempong, Wang, 2001). Many pediatric SCD treatment studies which utilize neuropsychological measures simply assess global IQ scores due to long-standing data on the relationships between IQ and functional outcomes, such as the likelihood of educational and vocational success. This methodology is flawed because measurement of specific executive processes is more informative than broad cognitive performance to understand SCD brain effects (Steen et al., 2003; Armstrong et al., 1996; Boni, Brown, Davis, Hsu, & Hopkins, 2001; Knight, Singhal, Thomas, & Serjeant, 1995). Treatment therapies also typically focus on the use of bone marrow transplantation (BMT), chronic transfusion therapy (CTT), and Hydroxyurea (HU) to treat symptoms associated with pediatric SCD.

At the present time, BMT is the only known curative therapy for SCD (Miller et al., 2000; Walters et al., 1996). The therapy consists of destroying SCD bone marrow and transplanting healthy bone marrow from a genetically-matched donor (Mehta, Afenyi-Annan, Byrns, & Lottenberg, 2006) Healthy bone marrow then produces normal hemoglobin. Shenoy (2011) reports that BMT procedures are under-utilized in the SCD population, referring to data that reveals that less than five hundred bone marrow

transplantation procedures were reported in the Center for International Blood and Marrow Transplant Research database. Although BMT is effective in curing SCD, the Seattle collaborative study found no change in neurologic complications of patients with SCD after BMT (Walters et al., 1997). This study investigated the effects of BMT in 22 children with a mean age of 9 year 9 months who had symptomatic SCD. MRI/MRA findings suggested that 2 children evidenced new neurologic damage after BMT and there was no change in the neurologic status for the remaining individuals in the sample. The lack of effectiveness of BMT for neurologic outcomes may well be due to the timing of BMT, which often occurs after extensive artery damage has already occurred. Furthermore, we do not yet know the long term effects of BMT in pediatric SCD (Walters et al., 2000). The gap in this literature suggests the importance of conducting more comprehensive neurological assessment before and after BMT treatment. Including neuropsychological measurement in treatment studies could provide more sensitive assessment of neurologic status after treatment, that is not evident by brain imaging or IQ scores alone.

CTT has been shown to be the most effective treatment in the prevention of recurring vasoocclusive episodes for children with SCD (Piomelli et al., 1985; Pegelow et al., 1995). These episodes occur when the circulation of blood vessels is obstructed by "sickled" red blood cells, causing neurologic injuries. CTT prevents these episodes by increasing brain oxygenation levels, suppressing HbS synthesis, and reducing hemolysis (Raj et al., 2004). The Stroke Prevention Trial in Sickle Cell Anemia (STOP) was pivotal in demonstrating the benefits of CTT for stroke risk in children with SCD. Findings indicated that children who were screened with TCD and subsequently received CTT

demonstrated less recurrence of overt stroke. This intervention also prevented first episodes of stroke for children with abnormal cerebral blood flow (Adams, McKie, Hsu, et al.1998). Likewise, Wilimas and associates (1980) used routine neuropsychologic testing along with radiologic and electrophysiologic studies to evaluate children with SCD and previous stroke who were treated with CTT for a 3 year period. Results revealed that their WISC -Revised mean FIQ increased from 63.7 to 67.2 with use of CCT. Presently, The Silent Infarct Multicenter Transfusion Trial is evaluating global IQ in those receiving blood transfusion compared to a comparison group of participants (study in progress). As noted by Schatz and associates (2002), use of broad scale IQ is limiting when compared to information obtained from specific measurement of cognitive functioning. More information about the treatment of CTT and cognitive effects could be understood if studies of this nature included measurements of narrow cognitive ability within the study design. Currently the literature lacks the data describing CTT and narrow range cognitive abilities due to the omission of neuropsychological measurement for these specific abilities. Without this information, treatment trials for SCD may miss important information about selective effects of treatment (e.g., benefits to specific cognitive abilities and associated neural systems) which could further inform brain imaging variables to study. The inclusion of modality specific cognitive measures within these studies could help to determine if the brain health benefits are global or specific with CTT.

HU (hydroxycarbamide), is a novel and effective treatment which increases fetal hemoglobin in patients with SCD, which inhibits the sickling of S-type hemoglobin. It increases the amount of oxygen distributed in the body and thereby decreases the chances

of vasoocclusion (Puffer, Schatz, Roberts, 2007). In a randomized control trial, HU was shown to reduce the number of hospitalizations and make pain crises milder by decreasing vasoocclusive complications and pulmonary tissue damage from acute chest syndromes (Charache et al., 1995) Subsequent studies have further supported the health benefits of HU therapy for the SCD population (Strouse et al., 2008; Lankron et al 2008). To date there is only one study that investigates the effects of HU on cognition in SCD (Puffer, Schatz, Roberts, 2007). This study examined 15 children with SCD on HU therapy compared to 50 healthy peers, statistically controlling for social/demographic factors in a retrospective, correlational design. Results indicated that children with SCD on HU performed better on verbal comprehension, fluid reasoning, and general cognitive ability tasks than control subjects. The findings from this study provide support for the cognitive benefits of HU therapy in pediatric SCD. Additional prospective investigations of HU and cognition should be conducted for more thorough understanding of treatment effects.

Summary

SCD refers to a category of inherited blood disorders which affect an estimated 1 out of 500 African Americans in the US (Charache, Lubin, & Reid, 1992). SCD complications can influence the neuropsychological functioning of the child. The cellular dysfunction associated with SCD can lead to a wide range of insults to the brain (Huttenlocher et al., 1984). This is evidenced by findings which suggest that the neurologic risk of stroke for children with SCD is more than 200 times the general population of healthy peers and the risk doubles for silent cerebral infarct (Early et al., 1998). Findings indicate overt stroke and SCD result in decreased cognitive ability

greater than silent stroke or SCD status (Steen et al., 2005; Schatz, Finke, & Roberts, 2004; Schatz, Finke, Kellett, & Kramer, 2002; Noll et al., 2001; Schatz et al., 2001). Along with decreased cognitive functioning in patients with SCD and stroke, children with SCD and silent cerebral infarcts show decrements in cognitive ability greater than SCD status but less than SCD and overt stroke (Bernaudin, Verlhac, Freard, Roudot-Thoraval, Benkerrou, et al., 2000; Schatz et al., 2001; Schatz & Buzan, 2006; Steen et al., 2005). Unfortunately cognitive deficits have also been found for children with SCD who have no history of brain abnormalities evidenced on MRI (Schatz et al., 2002; Steen 2003; 2005). Further understanding these cognitive dysfunctions are important for children with SCD because knowledge can help to identify disease related impacts as well as to inform treatment intervention research.

Treatment therapies such as BMT, CTT, and HU therapies have all been shown to significantly improve disease related effects of SCD (Walters et al., 1996; Adams et al., 1998; Charache et al., 1995). Although some treatment studies have utilized brain imaging to recognize neurologic treatment effects, only the Puffer and colleagues (2007) study, to date, has included neuropsychological assessment measures which assess narrow cognitive abilities. It is essential to add comprehensive neuropsychological measures to these studies so that the impact of treatment therapies on specific EF is understood. It is possible that treatments for SCD may differentially affect cognitive processes such that there is improvement in some areas yet greater deficit in other domains of EF. Furthermore, the absence of this information may cause treatment interventions to lack necessary focus to target specific cognitive systems.

Section II: Executive Function (EF)

Cognitive abilities such as those impacted in SCD are often described as EF. EF has been described in many different manners throughout its evolution and currently has no formal, broadly agreed upon definition. Historically, the EF concept largely grew out of awareness that IQ tests failed to capture important, higher-level cognitive abilities for everyday problem solving that could be seen in some individuals with neurologic disease. These cases often involved injury to regions of the prefrontal cortex. The concept of EF has often been used synonymously with the functions of prefrontal brain regions. More recently, scientists have begun defining executive function as also including important emotional processing and motivational/reward functions that influence judgment and decision making (sometimes called "hot" EF); however, classically EF has focused on "cold" cognitive processes involved in dealing with complex cognitive tasks. For this review we will focus on "cold" EF.

Rabbitt (1997) described EF as a portrayal of a group of cognitive actions such as: dealing with novelty, planning and implementing strategies for performance, monitoring performance, using feedback to adjust future responding, vigilance, and inhibiting task-irrelevant information. EF has also been defined as the cognitive processes which control and integrate component neurologic activities in support of adaptive behavior (Woodruff-Pak, 1997). Similarly, Zelazo and colleagues (1997) conceptualize EF as a macrostructure with executive sub functions which work together to accomplish the higher-order function of solving problems. For the purposes of this investigation, the EF definition developed by Gioia, Isquith, and Guy (2001) will be used to describe these

broad and narrow cognitive processes. Their definition describes EF as an umbrella expression that incorporates a compilation of inter-related processes which are responsible for purposeful, goal-directed behavior.

Theoretical models of EF are complex and have evolved greatly over the past 30 years. The following overview is meant to provide a sampling of the range of EF theories, rather than an exhaustive account, which is beyond the scope of the current review. In the early 1970s the EF concept emerged in the literature when Broadbent developed one of the first theories invoking the concept of EF. He reported that there are two separate mental processes, automatic and controlled (Shiffrin & Schneider, 1977). Similarly, Baddeley proposed a working memory model which contained the term "central executive" component of cognition (Baddeley & Hitch, 1974). After that, Posner further elucidated on the EF concept using the term "cognitive control" and created the foundation for much recent EF research (Posner & Synder, 1975). Posner theorized that focused attention on specific aspects of the environment is controlled by a distinct "executive" branch of the attentional system (Posner & Petersen, 1990). Additionally, Lezak (1983) developed an EF definition by explaining EF as the dimension of human behavior that deals with “how” behavior is expressed. It is also believed that the prefrontal cortex of the brain may mediate many component EF abilities; however the tasks controlled by these brain regions are numerous (Stuss et al., 2002). Stuss and colleagues (2002) has elaborated on critical EF functions mediated by the prefontral regions such as working memory, preparatory set, and inhibitory control, which he argues sub-erve the goal of integrating behavior over time (e.g., maintaining goals over time,

anticipating events). These classic theories led to the development of the current working models of EF.

More recent theories of EF include the Top-Down Inhibitory Control Model, Brown's EF Model, Miyake's "unity and diversity" theory, and the Anderson EF Model. The Top-Down Inhibitory Control Model focuses on the inhibitory control mechanisms involved in response control, memory, selective attention, theory of mind, and emotion regulation as the central mechanism to EF (Anderson & Green, 2001; Tipper, 2001; Stone & Gerrans, 2006). However, one of the main criticisms of this hypothesis is that it excludes the possible impacts of bottom up inhibition (Aron, 2007). Brown's model focuses on distinct component areas of EF. The six clusters of cognitive function in the Brown EF model includes activation, focus, effort, emotion, memory, and action are explained as dynamic and shifting functions (Brown, 2005). These models also emphasize domains of EF that often do not account well for EF, in an ecological sense. Most everyday tasks involving EF require the use of several of these domains. Miyake's theory identifies three main component areas for EF (inhibition, working memory, set shifting). However, Miyake theorizes that in tasks with functional significance all three areas are required to work in concert; thus, EF measures should tap into all three of these domains to be relevant to everyday behavior. Anderson's model focused more on integrating information from developmental and clinical literatures and identifies four main areas of EF with component skills that make up each area. Each domain is believed to be somewhat distinct (e.g., domains may be dissociated across different clinical conditions), yet contains component skills that could be measured in isolation.

Consequently, Anderson's theory falls somewhere between Brown and Miyake's models in terms of emphasizing the unity versus the diversity of cognitive skills in EF.

Summary

The EF field has progressed in both definition and theory in the past three decades. Currently there is not a standard definition of EF. However, the Gioia, Isquith, and Guy (2001) description of EF was utilized for this review. This definition was used because it allows for the study of component processes, which may inform us about the function of component neural systems. In addition, this definition upholds the notion that EF is critically important because of its role in adaptive behavior. Varied definitions of EF and evolving theories have negatively impacted the creation of measurement tools. Due to the wide ranging definitions and theories, EF tasks vary greatly in their format, content, and measurement strategy. It has often been the case that any task sensitive to prefrontal cortex injury will be described as an EF measure. A more detailed discussion of EF tools will be presented next.

EF Measurement

The lack of a recognized definition of EF makes the accurate assessment of cognitive abilities challenging. First, when invoking the EF construct, it is important to state the conceptual approach one is using. Such definitional statements are often lacking the literature. Many historically important measures of EF, such as the Wisconsin Card Sorting Test and the Category Test involved complex, cognitively heterogeneous, novel tasks. Although these measures are sensitive to brain dysfunction, these instruments often do not do a good job of identifying distinct neural systems that are affected nor do they perform well at identifying specific cognitive processes are affected by brain dysfunction.

Specific EF abilities have been shown to be related to specific neural systems. For example, the Wager and Smith (2003) meta-analysis found results indicated that that different EF processes are associated with specific cerebral areas. For instance, they found results which showed that the right inferior prefrontal cortex is frequently activated when it is necessary to manipulate information to perform a dual task (i.e., divided attention or shifting mental sets across tasks). Additionally, it was demonstrated that the superior frontal cortex shows activation once information is presented that has to be constantly restructured and when it is necessary that memory for temporal order is sustained (i.e., working memory). Due to neurologic associations such as these, many current approaches to identify EF capacities and processes use a task-based assessment of specific processes that have distinct neural associates (Hughes & Graham, 2002). Conceptually, this is consistent with the "diversity" approach of considering distinct, component EF processes.

Contemporary EF Measurement

The present day EF measurement approach involves an analysis of test psychometrics prior to administration, just as with any other psychological construct. This allows for a more thorough understanding of examinee performance. Although this may seem obvious to psychologists today, many neuropsychological tasks of EF were developed from behavioral neurology assessment tasks that lacked strong psychometric evaluation. Appropriate psychometric analysis includes an examination of considerations of the task requirements (content coverage, practicality of the methods), reliability, validity, and suitability of the standardized sample chosen to assess an individual (Groth-Marnat, 2003). These considerations are of particular importance for the SCD population

due to the uniqueness of disease related morbidity. Knowledge of test psychometrics aids in the appropriate interpretation of performance, given the characteristics of the SCD population and possible limitations of the measures.

Practicality of administration is an important assessment factor to examine prior to test administration. Recognition of how the measure approaches the EF construct is meaningful because the test may not be measuring one of the aspects of EF which are of interest for this population (Haynes, Richard, & Kubany, 1995). In addition, the context and manner in which the test is given is of particular importance. Differential ethnic group status between the examiner and examinee has been shown to impact the intellectual performance of children (Groth-Marnat, 1997; Sattler, 1992). Similarly Terrell, Terrell, & Taylor, (1980) found results which suggest that African American children obtained significantly higher scores on the WISC-R when administered the measure by an examiner of a similar race. Performance can be affected by educational attainment, as well. If the reading level of the test is higher than the subject's ability, scores may be adversely impacted.

Children with SCD often have poor education attainment and academic achievement due to school absences which result in poorer reading scores than healthy peers (Schatz et al., 2001; King et al., 2006). Additionally, test length may negatively impact performance on neuropsychological measures due to examinee fatigue or frustration, which may not be the constructs of interest (Groth-Marnat, 2003). Many neuropsychological test batteries are extensive and require sustained attention over long periods of time. The performance of children with SCD on these measures may be especially impacted by test length because disease related deficits have been shown to

impact sustained attention (White, Moinuddin, McKinstry, Noetzel, Armstrong, & DeBaun, 2006; Schatz et al., 2001).

Another psychometric issue especially related to pediatric SCD is the adequacy of the norm referenced comparative sample. Norm referenced developmental samples have become the typical evaluation method for the performance of individuals on neuropsychological measures (Sparrow et al., 2000). Norms in psychological assessment are described as the mean and standard deviation score among a category of individuals (McIntire & Miller, 2007). Individual test results are evaluated relative to psychometric norms. Standardized norm-referenced measures with dissimilar demographics from SCD characteristics may not adequately reflect performance in the pediatric SCD population. Generally, the composition of normative groups is stratified by age and/or education level (Cicchetti, 1994). Children with SCD may not meet usual age range expectations due to disease related morbidity (Stevens et al., 1986; Zemel et al., 2007). Furthermore, most EF tasks have restricted age/ability ranges in contrast to SCD treatment studies (Gaston et al., 1986; Ballas et al., 2000). These studies often use broad age range convenience samples, the implications of this dissonance are possible increases in errors related to interpretation. Additionally, measurement precision is often not the same across different ability levels (McIntire & Miller, 2007). Finally, test generalizability is limited to the characteristics of the norm sample. Characteristics of the norm sample should be investigated when compared to the examinee, particularly in the SCD population in part because health status has been shown to affect test psychometric scores (Johnson, 2006). The performance of SCD children should be interpreted with caution if the characteristics of the SCD sample are extremely divergent from demographics of the normative sample.

Methods of Isolating Specific EF Processes

Modern day neuropsychological measures are founded on traditional psychometric theory. The traditional model often compares an overall raw score of an individual on a cognitive task to norm referenced scores of the standardized sample (McGrew & Flanagan, 1998). This comparison is hypothesized to allow inference regarding individual performance to a control group. Ideally, this approach provides a fair assessment of current cognitive function (Lezak, Howieson, Loring, Hannay, & Fischer, 2004). As noted previously, similarities between the norm group and the individual may influence performance on a measure (Johnson, 2006). Most commonly, cognitive performance evidenced by classic cognitive tests such as the WISC are evaluated using this method. This has sometimes been described as an "achievement" approach to test evaluation because it reflects the person's capability (or lack of capability), but often provides little information about why or how the person achieved a particular score.

A theoretical alternative to the traditional model for identifying neuropsychological ability is the subtractive factors method. Donders (1868/1869) developed a reaction time methodological paradigm to evaluate the time involved in a set of specific sequences between the presentation of a stimulus event and the initiation of response (Schweickert, 1978). The subtractive factors approach hypothesizes that the time it takes to complete a particular process can be estimated by adding that process to a task and taking the difference in reaction time between the two tasks (Donders, 1969). This method has three components and has been used for recognition, discrimination, and response selection measurement (Gottsdanker & Shragg, 1985). Critics of this theory challenged the central underlying assumption that one mental process can be added to a

task without affecting the time to complete other mental processes (Ulrich, Mattes, & Miller, 1999). Currently the empirical research of Cattell support this criticism, nevertheless Cattell did not analyze this hypothesis within Donder's formal model (Cattell, 1886; 1888; 1947). Thus, although the subtractive factors approach may not work with all cognitive tasks, it represents a potentially powerful method to assess component skills that make up a more complex task. This approach may be well suited to the measurement of EF, as we may be interested in both overall performance level as well as measuring discrete cognitive skills. Subtractive factors methods have recently been applied to norm referenced cognitive ability measures such as Delis–Kaplan Executive Function System (D–KEFS), NEPSY - Second Edition (NEPSY–II), California Verbal Learning Test (CVLT–II) and Wide Range Assessment of. Memory and Learning - Second Edition (WRAML).

Another alternative is the additive factors approach. The additive factors method is an additional psychometric approach for the isolation of cognitive processes (Sternberg, 1969). Sternberg further developed Donder's theoretical paradigm to create this model. Sternberg's additive factors logic interpretation of reaction time pattern does not require the assumption of discrete stages, like subtractive factors, as long as it is assumed that the final output of a stage does not vary as a function of the manipulations. The central assumption of the additive factors method is the processing sequence between stimulus and response that consists in a series of discrete stages. It has been shown to be a powerful tool for studying reading process models (Raake, 2006; Massaro, 1975). However, the additive factors approach is currently not used in any norm-referenced tests. This is likely because the method requires a thorough measurement of at least two

factors over multiple levels of those factors, including their interaction. Thus, an additive factors approach would add to the length and burden of many tasks. It is therefore best suited to narrow assessment of a few discrete areas cognitive ability as opposed to the broader assessment strategy used in most clinical contexts.

EF Psychometrics

There are several different types of test validity and reliability essential to the psychometric properties of psychological tests. Assessment tools designed to report the objective performance of an individual's neuropsychological functioning have traditionally been more useful to treatment research than subjective observations and interviews due to their greater sensitivity to change (Groth-Marnat, 1999). Once a test has been shown to maintain similar performance between testing sessions, additional psychometric qualities such as reliability and validity should be evaluated.

Reliability assesses the extent to which a construct is consistently measured within individuals (Loevinger, 1957; Clark & Watson, 1995). There are several common classes of reliability which include inter-rater, intra-rater, test-retest, and internal consistency (Groth-Marnat, 1990). For the purposes of this review, test retest and internal consistency reliabilities will be described for EF measures utilized in pediatric SCD literature. Test retest reliability is determined by administrating the test and then repeating test administration during another testing session. The second testing session must occur soon enough that changes in the underlying construct would not be expected (Meeker and Escobar, 1998). The test retest coefficient is calculated by a correlation of two different scores, on two different occasions, by the same individual (Cortina, 1993; Iacobucci & Duhachek, 2003). When the purpose of the measure is to make predictions

and ensure that stable constructs do not fluctuate highly, the evaluation of test retest reliability is recommended (Groth-Marnet, 1990). If test retest reliability is not known prior to administration of the measure, unreliable results could be obtained and false inferences could be made when the measure is administered before and after treatment of SCD.

Alternatively, internal consistency assesses the uniformity of test items across the measure and is the correlation between different items on the same test (Cortina, 1993). The identification of internal consistency is valuable on measures with longer formats and various subtests or sub scales (Groth-Marnet, 1990). These forms of reliability were chosen because they are the best forms of reliability to understand both the nature of the EF and to inform treatment evaluation. Internal consistency is essential to the analysis of measures in the pediatric SCD population because brief measures may be chosen due to neurologic effects of the disease (e.g. decreased sustained attention). However, the shorter format may not adequately represent the EF construct of interest.

Validity is one of the most critical factors in psychometric test construction. Validity refers to the amount of which psychological tools assess the construct that they purport to measure (Guion, 1980). Content, criterion, and construct validity are the broad categories of test validity.

Content validity requires a degree of agreement to define a particular construct. The meaning used to assess performance can affect the inferences that may be drawn from the obtained data (Groth-Marnet, 1990). Depending on the definition utilized, it may over or under represent, omit facets, or reflect factors on the periphery of the EF construct (Haynes et al., 1995).

The criterion validity of EF measures used in the pediatric SCD population was evaluated for this review. This factor was chosen because EF is a complex multi-dimensional construct and opinions about its characterization and evaluation are often subjective and divergent. Criterion validity is the degree to which a test performs as expected based on an external standard (or criterion). In many instances researchers use another established assessment test which shares a similar theoretical orientation as a criterion measure (similar to convergent validity). This psychometric was chosen as an appropriate representation of validity for measures used in this population. The simplest way to understand criterion validity is to compare a psychological measure it to a gold standard test which measures the same construct. Sensitivity and specificity for clinical outcomes also inform whether a measure is highly validated or not and is sometimes used as an external criterion measure (White et al., 2006). In this context, sensitivity illustrates the accurate identification of children with known neurologic disease or suspected to be at high risk for neurologic disease within the SCD population. In contrast, specificity demonstrates the accuracy involved in correctly identifying children without neurologic disease.

Construct validity is based on the extent to which the measure correlates (positively or negatively) with other measures as expected from one's theoretical model or known empirical relationships. Thus, a measure of working memory should correlate with other measures of working memory, other tasks known to be associated with working memory ability, but not with unrelated constructs (e.g., positive affect). In batteries of tasks, the correlation coefficient is calculated for each measure and sub scales. Inter-correlations are then determined and compared to estimate the degree to

which the two measures relate to each other. Comparisons among neuropsychological tests, which measure EF, are informative to the evaluation of cognitive performance across measures. These comparisons are also used in the identification of the most appropriate measures to use for individuals diagnosed with pediatric SCD. Another approach to construct validity is to use factor analytic methods to determine how well the measure conforms to the theory or framework used to develop the measure. Traditionally, construct validity has received the most attention in the development of norm referenced measures and test developers can demonstrate at least moderate to high construct validity. Thus, though construct validity is critical, criterion validity is the more neglected type of validity and the type of validity that maps on better to clinical decision making.

Summary

Quality procedures for the selection of psychometric measures should be conducted in general, however they should be especially considered when administering these measures to a chronically ill pediatric population. Method of test interpretation, normative sample quality, reliability, and validity all influence the appropriateness of cognitive measures selected for use in pediatric SCD. The educational level, manner in which the test is administered, social-interactional context, and length of assessment have all been shown to impact the performance of individuals on a measure (Groth-Marnet, 1990). Psychometricians should demonstrate culturally sensitive behaviors, during test administration, towards the reduction of possible contextual effects. Additionally, brief tests with lower reading levels given in a standardized manner are well suited for this population given their neurologic considerations.

Similarity between the norm group and the individual may also affect the choice of assessment test. Dissimilarity between norm group and examinee has been shown to impact the examinee's performance reported by a neuropsychological measure (Johnson, 2006). The norm sample used in a neuropsychological assessment for the purposes of comparing children with SCD should include a stratified sample. This sample should contain a census or over representation of African Americans children with varied general health status. Additionally, the subtractive factors approach to isolating specific abilities may be best suited for use with pediatric SCD. This approach may be superior due to the within subject methodology versus drawing inferences for the comparison of a dissimilar normative group (Gottsdanker & Shragg, 1985). Furthermore, the reliability and validity of a neuropsychological measure should be evaluated prior to use (Groth-Marnet, 2000). In particular the level of test retest reliability, internal consistency, and criterion validity should be evaluated and well understood when used with children diagnosed with SCD.

Criticisms of Cognitive Measurement

There are several different criticisms for the use of EF measures in specialized populations. One criticism to utilizing these measures in treatment outcomes is the lack of information published regarding practice effects (Filskov & Boll, 1981). Practice effects take place when a child performs a task and then performs it again on another occasion. They are recognized to exist for neuropsychological tests, but the magnitude of these effects is not well known for specific measures (Strauss, Sherman, & Spreen, 2006). In treatment studies with pre-post only designs, this can pose problems for data interpretation. Difficulties can occur such as the child's scores during the second

administration improve due to previous exposure to the test, not to the treatment therapy. Although the magnitude of practice effects are not well understood; thoughtful review of the type of measure, test retest reliability, age range, and characteristics of the subjects can greatly reduce the error variability associated with cognitive measurement (Dikmen et al., 2000).

Another critique for the use of neuropsychological measures in treatment studies is the association between cognitive measures and actual daily functioning. A strong correlation between the test scores and functioning is preferred because it validates the use of the test as a predictor of behavior. Hairman (1991) found results that indicate a modest degree of association between cognitive effects and functional effects in children with SCD and stroke. However, the Hairman (1991) study included only children with stroke and may have suffered of a restriction of range for functional behavioral outcomes. Data has demonstrated, more with adults than children, that cognitive performance on neuropsychological measures is moderate in relation to measures of activities of daily living (Farias, Harrell, Neumann, & Houtz, 2003; Carter, Oliviera, Duponte, & Lynch, 1988). In contrast, cognitive measures have shown large effect sizes for academic achievement and educational attainment in the pediatric SCD literature (Schatz 2004, Schatz et al., 2001; Nettles, 1994).

There are also critiques of neuropsychological assessment which relate directly to the multifaceted risk factors often found in SCD. Criticisms include the lack of cognitive score interpretation in light of broader psychosocial factors (Strauss, Sherman, & Spreen, 2006). In particular, factors such as parent-family factors, educational experiences, and motivational factors for test performance may all be relevant to test outcomes (Thompson

et al., 2004). In terms of social and family factors, Thompson and colleagues (2004) found results that indicate one fourth of the parents of a child with SCD experienced clinically significant levels of psychological distress and poorer parent adjustment evidenced high levels of daily stress, less knowledge about child development, and lower expectations of efficacy. In addition they found results which indicated that poorer cognitive functioning was associated with the learned-helplessness attributional style for parents of children with SCD. This psychosocial interaction may pose a problem for isolating disease-specific effects or treatment-specific effects (Hertzog & Nesselroade, 2003). Interpretation of cognitive scores in view of the psychosocial factors is best practices for psychological assessment and should be utilized with the pediatric SCD population (Sue et al., 1985).

Specific issues in the selection of measures for the SCD population also include test fairness for African Americans and lower SES children (Schatz et al., 2004). The SCD population is the US contains an overrepresentation of African Americans (Wang et al., 2001). The ethnic minority composition of the SCD population should be considered prior to test selection. Examination of the cognitive performance for African Americans has been conducted extensively and recommendations have been suggested which promote fairness in neuropsychological testing (Unversagt, et al., 1996; Diehr, Heaton, Miller, & Grant, 1998; Gladsjo et al., 1999). An example of recommendations that have been made for the construction of culturally fair neuropsychological test are is examination of "cultural contamination" prior to making predictive conclusions about the examinee (Helms, 2002; Darlington, 1971). Helms (2002) describe cultural contamination as a significant correlation between a test score and a total score on a

racial identity questionnaire after controlling for the criterion variable. Relatively few neuropsychological measures utilize this recommendation, and there are no measures identified in the pediatric SCD literature which follow this guideline.

To address concerns regarding minority performance on psychological measures, the Standards for Educational and Psychological Testing, published jointly by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA & APA) in1999 recommend evaluating neuropsychological measures in terms of fairness. Fairness is estimated by four principal methods which include instrument bias, equitable treatment, measurement equivalence, and diagnostic validity (AERA & APA, 1999). To date only one measure, D–KEFS, has been found to adhere to fairness standards in the pediatric neuropsychological SCD literature. Selection of neuropsychological measures which have been evaluated by the fairness guidelines may help to ensure psychological measurement equity across culture for African American children which are diagnosed with SCD.

Understanding the multifaceted nature of SCD is vital to the selection of neuropsychological tests that are most appropriate for this population. SCD affects the general health as well as the neurologic condition of children with SCD. Issues such as such as fatigue from anemia, CVA, silent cerebral infarct, increased blood flow velocities may pose a significant confound for test score interpretation (Palermo, Schwartz, Drotar, & McGowan, 2002). When evaluating measurement tools for use with the pediatric SCD population, health effects may impact performance unrelated to the cognitive instrument. For example, children with SCD are often in pain related to vasoocculsive crisis that do

not require hospitalization (Platt, Thorington, Brambilla, et al. 1991; Vichinsky, Johnson, Lubin, 1982; Shapiro, Dinges, Carota-Orne, Ohene-Frempong, Orne, 1990). An examiner unfamiliar to the effects of pediatric SCD may not inquire as to the child's current pain level. The absence of the information may lead to inaccurate results which are associated with pain instead of cognition. Therefore, administering a neuropsychological test to a child suffering from an unreported low level pain crisis may increase measurement error in the cognitive score. It is important to comprehend all of the properties related to a psychological measure that are not the target of study, but nonetheless result in group differences in test scores. Individual conditions such as health status may impact the fairness of administration (Van de Vijver and Poortinga, 2005).

Summary

The advantages associated with the administration of neuropsychological assessments for children with SCD are noteworthy. Strengths associated with cognitive measurement for pediatric SCD include the identification of specific cognitive abilities in SCD treatment studies. Likewise, a more thorough understanding of brain mechanisms which affect cognitive processes can be discovered for individuals with pediatric SCD. The addition of neuropsychological measures to the existing body of literature will help to propel SCD treatment to a greater comprehensive knowledge base regarding both treatment and disease related effects.

Disadvantages related to using cognitive measures with this population are various. Practice effects are an instance of one such difficulty. The magnitude of practice effect for specific measures is unknown (Strauss, Sherman, Spreen, 2006). Additionally, the degree of association between cognition and functional living, primarily with adults,

has also been a critique of including cognitive measures to treatment studies (Hairman, 1991). Furthermore, psychosocial factors which influence disease-specific or treatment-specific effects may impact performance on neuropsychological measures (Hertzog & Nesselroade, 2003). Generally neuropsychological measures have been criticized for their fairness and disparate impact on African Americans which is salient due to the composition of the SCD community (Wang et al., 2001). Finally, the health related morbidity for children with SCD may impact performance unrelated to cognitive scores (Platt, Thorington, Brambilla, et al. 1991). Given the appropriate considerations, each of these potential confounds can be mitigated thereby allowing for the correct interpretation of EF performance.

Section III: Existing Literature on Cognition in Pediatric SCD

A review of the pediatric SCD literature revealed a total of 40 studies which contained cognitive measures to assess neuropsychological functioning, see Table 1.1. This table demonstrates the wide variation of neuropsychology measures that are used across the pediatric SCD literature. This lack in consistency has contributed to the complexities associated with understanding the cognitive profile of children with SCD. Additionally, this disparity in the literature has indirectly promoted the lack of narrow EF measurement in the treatment literature. These tests were selected by researchers to identify cognitive abilities in the pediatric SCD literature. The measurement batteries include six versions of the Wechsler Intelligence, three editions of the Woodcock Johnson Cognitive tests, DAS, Kaufman Assessment Battery for Children (KABC), D–KEFS, Denver Developmental Screening Test, Developmental Neuropsychological

Assessment, Luria-Nebraska Neuropsychological Battery, Stanford-Binet, and the Behavior Rating Inventory of Executive Function (BRIEF).

Among the 18 neuropsychological batteries utilized; the Wechsler, Woodcock Johnson, and the DAS scales emerged as the most commonly administered in the pediatric SCD population. Many of these studies analyzed various facets of cognitive ability with the Wechsler Scales. Although this measure is commonly used in the literature, Wechsler (1991) reports that these scales measure only a few factors of cognition among the numerous that have been proposed within the Cattell-Horn-Carroll (CHC) theory. The CHC theory identifies 10 broad and over 70 narrow abilities to inform a distinct EF profile (Flanagan & Harrison, 2005). The use of this scale may therefore reveal an incomplete picture of cognition for this population. In contrast to the Wechsler scales, the WJ-III is reported to measure a multitude of factors within the CHC framework and was selected for use in 12 pediatric SCD studies (Woodcock, McGrew, & Mather, 2001). The DAS was also used, in six studies, to further understand EF in this population. This cognitive battery was not developed from any single theory of cognitive ability but instead derives from sophisticated statistical procedures. Additionally, various neuropsychological measures designed to specifically evaluate narrow aspects of EF were also used to comprehend the nature of cognition in children diagnosed with SCD. Table 1.1 reveals the administration of 58 different cognitive tests for an analysis of EF to the pediatric SCD population. These cognitive measures are assorted for 2 primary reasons: (1) the multifaceted definition of EF and (2) the lack of standard assessment guidelines for this population.

General descriptive information for the selected studies was also examined, see Table 1.2. The average number of participants within the 30 studies included in this review was 163.8 (*SD* = 423.9) participants with a mean age of 10.6 (*SD* = 2.4) years. However, this statistic is misleading because 84% of the studies had samples less than 100. By excluding Steen and colleagues (2005), which had an outlier sample of 2254, the mean participant pool was 82.5 (*SD* = 71.9) with a mean age of 10.8 (*SD* = 2.3). Sickle Cell Anemia (HbSS) was the predominant subtype of SCD in most studies. Moreover, this table clearly reveals the varied age range (1 year 8 months – 16 years nine months) and dissimilar methodological considerations in the pediatric SCD literature. An illustration of this these practices is evidenced when analyzing the use of the Wide Range Achievement Test (WRAT) in pediatric SCD studies. The WRAT was used in 3 research studies (Cohen et al., 1994; Wasserman et al., 1991; Fowler & Whitt, 1988). Although this measure was used to investigate academic achievement in children with SCD, only 2 of the 3 studies were similar to one another in age and sample group characteristics. Children, mean age 9 years and 9 months, with SCD were grouped according to left versus right silent cerebral infarct in the Cohen and colleagues (1994) article. In contrast, Wasserman and colleagues (1991) compared children, with a mean age 11 years 5 months, with SCD and no stroke to siblings. Likewise, Fowler & Whitt (1998) also investigated children with SCD and no stroke compared to healthy peers, mean age 12 and 11 months. Comparing children of different age ranges, across the literature, may be problematic as this assumes any cognitive deficits are constant across different ages. There are also potential issues with cohort effects as the quality of medical care has changed over time.

It is also unmistakable that norm referenced assessment practices are varied across pediatric SCD studies. The psychometric norm characteristics for each neuropsychological measure in this literature were either compared to a standard or local normative group of children. As shown in Table 1.2, an estimated forty percent of pediatric SCD studies selected a combination of measures which include both local and standard norms to assess EF. Stewart and Kaminski (2002) suggested that local norms provide an increased advantage of providing meaningful information regarding performance for a particular demographic on psychological measures. These authors also report that local norms decrease the likelihood of bias because a child's test performance is compared to other children whose demographic factors are similar. Furthermore, they also demonstrate that local norms are useful in facilitating the identification of strengths and weaknesses. The majority of studies within pediatric SCD did not utilize measures which included local norms. However contrasting the cognitive results of individual children with SCD tested within these research studies to one another or similar group of children with SCD seems more informative than to a healthy group of nationally-representative children. Increased use of measures which include a normative sample more appropriate for children with SCD can help to reveal a more accurate picture of cognition is pediatric SCD.

The psychometric properties of 30 identified neuropsychological measures and relevant subtests; when reported in the pediatric SCD literature, was reviewed. Age range and method of test interpretation were included as an evaluation technique to understand the psychometric properties of each measure. Additionally, test retest and internal consistency were analyzed to understand the psychometric reliability of each measure.

Also, convergent validity was examined to assess psychometric validity across neuropsychological measures. The quality of the normative sample was reviewed with seven different criteria (see Table 1.4). The selected psychometric factors were reviewed to inform the appropriateness of test selection for children with SCD.

Neuropsychological measures administered for this population encompass several different age ranges. Overall, the age ranges of neuropsychological measures were appropriate chosen for the study samples. However the Wechsler Memory Scales (WMS) were utilized outside of the intended psychometric age range by Watkins and colleagues (1989). The WMS suggested for use with individuals age 16-89, nevertheless the authors utilized this measure to understand memory functions in a sample of children with SCD mean age 10 years 6 months. Results reported that are associated with this measure should be interpreted with caution due to the age range guidelines suggested with this instrument.

The method of test interpretation within neuropsychological measures for pediatric SCD is limited to the traditional and subtractive factors measurement approaches. As described earlier, subtractive factors may be the more potent method of test interpretation for children with SCD due to within subject comparison. However, it appears that only eight cognitive measures utilize this method of interpreting examinee performance among pediatric SCD studies (see Table 1.3). Specifically, use of these measures is generally accepted in research contexts; however due to their "experimental" nature, most data regarding psychometric properties is not published. It is clear that measures which utilize the traditional "achievement" approach of understanding cognition have dominated the pediatric SCD literature.

There was extreme variability in reliabilities evaluated within and across the 30 measures selected for use with the pediatric SCD population. Several works were consulted to describe statistically adequate measurement practices. Ultimately reliability evidence was determined by recommendations by Strauss and colleagues (2006). Test retest reliability coefficients were determined as: "Very High (≥.90), High (.80-.89), Adequate (.70-.79), Marginal (.60-.69), Low (.50-.59), Extremely Low (≤.49).'' Additionally, internal consistency reliabilities were set as: Excellent ($\alpha \geq .9$), Good ($.9 > \alpha \geq .8$), Acceptable ($.8 > \alpha \geq .7$), Questionable ($.7 > \alpha \geq .6$), Unacceptable ($.5 > \alpha$). In general reliability coefficients ranged within measures from extremely low to high, depending on the subscale of interest. For instance, The Self Ordered Pointing Test reliability associated with test retest was adequate, whereas the internal consistency was high. Additionally the CVLT–II, long delay recall subtest has marginal test retest reliability but excellent internal consistency. On average, when reported, test retest reliability for the selected measures across the studies is adequate and internal consistency is high. However, it should be noted, data was missing for many of the subscales utilized to asses cognitive functioning. An example of this missingness is evident in the Detroit Test Learning Aptitude 2nd ed. Reported test retest reliability varied from adequate to very high and internal consistency from acceptable to excellent on this instrument. However, specific subtest reliability coefficients for the measures shown to be sensitive to SCD could not be obtained. Ideally, reporting of reliabilities would be readily available and consistent with one another, such as in the case of the Decision Speed subtest of the WJ-III. The test retest reliability is high and the internal consistency is excellent.

Overall measures in this literature demonstrated a strong relationship to other cognitive measures which measure the same features of EF (see Table 1.3). The following convergent validity correlations established by Strauss et al. 2006 were utilized for this review: Strong (.50 to 1.0), Moderate (.30 to .49), Low (.10 to .29). A few neuropsychological measures diverge from the overall prevalence strong associations in the research literature. For instance, measures such as CVLT–II, Gordon Continuous Performance Task, Kagan Matching Familiar Figures Test, and Self Ordered Pointing Test indicated moderate associated with similar tests of cognition (see Table 1.3). These modest relationships are not surprising due to the nature of EF and measurement practices. Additionally, the Tower of Hanoi and Wisconsin Card Sorting Test have low associations, at best, with other measures of cognition.

Normative sample characteristics were evaluated by several factors to develop a general set of quality guidelines (see Table 1.4). The guidelines created for this review originated from recommendations for normative samples by Mitrushina and associates (2005). The 10 criteria chosen describe the degree of support for the normative samples utilized by each neuropsychological measure. Additionally, these guidelines include cultural considerations indicated by Suzuki and Ponterotto (2008) relevant to the SCD populations. Notably, evaluation of the quality criteria requires an extent of subjectivity in judging psychometric standards. Norm quality consists of an analysis of: sample size, representativeness of United States (US) census data, stratification, etlhnicity by stratification, geographic diversity, African American cultural effects, health condition, special education inclusion, and exclusion/inclusion criteria. An analysis of these factors

resulted in a categorization of deficits on the following continuum "excellent ≥ 8 factors, adequate ≥ 6 factors, good ≥ 4 factors, and poor ≤ 3."

Normative samples from 13 of the 30 articles analyzed were reviewed to be of "poor" quality (see Table 1.3). This status was the result of a combination of factors such as small sample size, no representativeness from US census data, no stratification, limited geographic diversity, strong African American cultural effects, etc. An example of a neuropsychological measure with a designation of "poor" for the normative sample is the Connors's Continuous Performance Test. The clinical sample was comprised of 670 individuals who were stratified by eight age bands. This sample was not representative of the US census and subjects were only selected from five US states and a province in Southern Ontario. The general health condition of the sample was not indicated and children receiving special education were excluded from the sample. However, the mental health status of the sample was reported and no ethnic/racial effects were reported as well. This sample lacks the necessary properties to appropriately compare the performance of children with SCD. A more appropriate measure, with a description of "excellent" for the test norms is the WJ-III. The sample size is large, consisting of 5972 children stratified by both individual and community socioeconomic variables. Additionally, the validation sample is geographically diverse, with individuals selected from over 100 geographic regions within 27 US states. Similarly, comparable scores have been evidenced for African American children when compared to their Caucasian peers. Albeit, it is significant to mention the general health conditions of the sample were not reported, the only exclusion criteria were less than one year of fluent English speaking.

Despite this, the test norms represent an acceptable comparative sample for children with SCD, which provide a more accurate representation of cognitive performance.

Norms, Measurement Approach, and Reliability of Cognitive Measures Used to Date

The psychometric properties of 30 identified neuropsychological measures and relevant subtests, when reported, in the pediatric SCD literature indicate many measures used to date have been lacking in one or more key dimensions. Age range was included in this range because it is desirable for many clinical studies to be able to include a wide range of ages. Method of test interpretation were included as an evaluation technique to understand the whether the measure allows for more fine-grained analysis of cognitive skills. Additionally, test retest and internal consistency were analyzed to understand the psychometric reliability of each measure. Convergent validity was examined to assess psychometric validity across neuropsychological measures. The quality of the normative sample was reviewed with ten different criteria. The selected psychometric factors were reviewed to inform the appropriateness of test selection for children with SCD (see Table 1.4).

Neuropsychological measures administered for this population encompass several different age ranges. Overall, the age ranges of neuropsychological measures were appropriate chosen for the study samples. However the WMS were utilized outside of the intended psychometric age range by Watkins and colleagues (1989). The WMS suggested for use with individuals age 16-89, nevertheless the authors utilized this measure to understand memory functions in a sample of children with SCD mean age 10 years 6 months. Results reported that are associated with this measure should be interpreted with caution due to the age range guidelines suggested with this instrument.

The method of test interpretation within neuropsychological measures for pediatric SCD is limited to the traditional and subtractive factors measurement approaches. As described earlier, subtractive factors may be the more ideal method of test interpretation for children with SCD due to within subject comparison to isolate specific skills. However, it appears that only eight cognitive measures utilize this method of interpreting examinee performance among pediatric SCD studies (see Table 1.4). For some of these measures using subtractive factors were "experimental" (not norm-referenced measures designed for individual-level interpretation in clinical contexts). Specifically, use of these measures is accepted in the study contexts; however due to their "experimental" nature most data regarding psychometric properties is not published. On the contrary, it is clear that measures which utilize the traditional "achievement" approach of understanding cognition have dominated the pediatric SCD literature.

There was extreme variability in reliabilities evaluated within and across the 30 measures selected for use with the pediatric SCD population. Several works were consulted to describe statistically adequate measurement practices. Ultimately reliability evidence was determined by recommendations by Strauss and colleagues (2006). Test retest reliability coefficients were determined as: "Very High (≥.90), High (.80-.89), Adequate (.70-.79), Marginal (.60-.69), Low (.50-.59), Extremely Low (≤.49).'' Additionally, internal consistency reliabilities were set as: Excellent ($\alpha \geq .9$), Good ($.9 > \alpha \geq .8$), Acceptable ($.8 > \alpha \geq .7$), Questionable ($.7 > \alpha \geq .6$), Unacceptable ($.5 > \alpha$). In general reliability coefficients ranged within measures from extremely low to high, depending on the subscale of interest. For instance, The Self Ordered Pointing Test reliability associated with test retest was adequate, whereas the internal consistency was

high. Additionally the CVLT–II, long delay recall subtest has marginal test retest reliability but excellent internal consistency. On average, when reported, test retest reliability for the selected measures across the studies is adequate and internal consistency is high. However, it should be noted, data was missing for many of the subscales utilized to asses cognitive functioning. An example of this missingness is evident in the Detroit Test Learning Aptitude 2nd ed. Reported test retest reliability varied from adequate to very high and internal consistency from acceptable to excellent on this instrument. However, specific subtest reliability coefficients for the measures shown to be sensitive to SCD could not be obtained. Ideally, reporting of reliabilities would be readily available and consistent with one another, such as in the case of the Decision Speed subtest of the WJ-III. The test retest reliability is high and the internal consistency is excellent.

Overall measures in this literature demonstrated a strong relationship to other cognitive measures which measure the same features of EF (see Table 1.3). The following convergent validity correlations established by Strauss et al. 2006 were utilized for this review: Strong (.50 to 1.0), Moderate (.30 to .49), Low (.10 to .29). A few neuropsychological measures diverge from the overall prevalence strong associations in the research literature. For instance, measures such as CVLT–II, Gordon Continuous Performance Task, Kagan Matching Familiar Figures Test, and Self Ordered Pointing Test indicated moderate associated with similar tests of cognition. These modest relationships are not surprising due to the nature of EF and measurement practices. Additionally, the Tower of Hanoi and Wisconsin Card Sorting Test have low associations, at best, with other measures of cognition. This finding is also not surprising

given the concurrent and complex cognitive functions that are necessary to complete these tasks.

Normative sample characteristics were evaluated by several factors to develop a general set of quality guidelines (see Table 1.4). The guidelines created for this review originated from recommendations for normative samples by Mitrushina and associates (2005). The 10 criteria chosen describe the degree of support for the normative samples utilized by each neuropsychological measure. Additionally, these guidelines include cultural considerations indicated by Suzuki and Ponterotto (2008) relevant to the SCD populations. Notably, evaluation of the quality criteria requires an extent of subjectivity in judging psychometric standards. Norm quality consists of an analysis of: sample size, representativeness of United States (US) census data, stratification, ethnicity by stratification, geographic diversity, African American cultural effects, health condition, special education inclusion, and exclusion/inclusion criteria. An analysis of these factors resulted in a categorization of deficits on the following continuum "excellent $\geq$ 8 factors, good $\geq$ 6 factors, adequate $\geq$ 4 factors, and poor $\leq$ 3."

Normative samples from 13 of 25 test analyzed were reviewed to be of "poor" quality (see Table 1.4). This status was the result of a combination of factors such as small sample size, no representativeness from US census data, no stratification, limited geographic diversity, strong African American cultural effects, etc. An example of a neuropsychological measure with a designation of "poor" for the normative sample is the Connors's Continuous Performance Test. The clinical sample was comprised of 670 individuals who were stratified by eight age bands. This sample was not representative of the US census and subjects were only selected from five US states and a province in

Southern Ontario. The general health condition of the sample was not indicated and children receiving special education were excluded from the sample. However, the mental health status of the sample was reported and no ethnic/racial effects were reported as well. This sample lacks the necessary properties to appropriately compare the performance of children with SCD. A more appropriate measure, with a description of "excellent" for the test norms is the WJ-III. The sample size is large, consisting of 5972 children stratified by both individual and community socioeconomic variables. Additionally, the validation sample is geographically diverse, with individuals selected from over 100 geographic regions within 27 US states. Similarly, comparable scores have been evidenced for African American children when compared to their Caucasian peers. Albeit, it is significant to mention the general health conditions of the sample were not reported, the only exclusion criteria were less than one year of fluent English speaking. Despite this, the test norms represent an acceptable comparative sample for children with SCD, which provide a more accurate representation of cognitive performance.

Cognitive Measures and SCD Morbidity (Criterion Validity)

Neurologic damage caused from SCD related mechanisms often result in stroke, silent infarcts, and cerebral blood flow abnormalities. In addition, general disease severity factors such as anemia may also be related to brain health. Data were reviewed to identify neuropsychological test sensitivity associated with different types of SCD disease effects.

The Stroke Prevention Trial in Sickle Cell Anemia identified the importance of TCD intervention for children identified at high stroke risk. Three research studies have been identified that address neuropsychological sensitivity for elevated TCD velocities and performance on the Wechsler scales (Bernaudin et al., 2000; Kral et al., 2003; Kral &

Brown, 2004). Bernaudin and collegues (2000) conducted an investigation which examined the first study evidencing an effect of thrombocytosis and revealing that silent stroke alone is not a factor of cognitive deficit when not associated with low hematocrit or thrombocytosis. Results indicated cognitive deficits were observed in SCD patients with a history of overt stroke and that cognitive deficits are apparent for children with SCD and silent stroke evidenced by infarcts on MRI, severe chronic anemia, and elevated TCD values. Another research team found that children diagnosed with SCD and abnormal TCD values evidenced greater executive dysfunction than children with SCD and conditional TCD values (Kral et al., 2003). It also appears that children with abnormal TCD values and SCD performed differentially worse than children with conditional and normal TCD values (Kral & Brown, 2004). Results from these studies suggest that Wechsler scales are sensitive to cognitive deficits in children with SCD and elevated TCD values.

Similarly, SCD researchers have begun to investigate the utility of the BRIEF in evaluating cognitive functioning behaviors and SCD related neural mechanisms in the pediatric SCD population. An early investigation of pediatric SCD and TCD revealed that children diagnosed with SCD with normal TCD findings demonstrated better executive functioning than children with SCD and conditional or abnormal TCD values (Kral, Brown, Nietert, Abboud, Jackson, and Hynd, 2003). Kral and colleagues (2004) also found that teacher ratings of children with SCD and abnormal TCD demonstrate poorer cognitive functioning than children with SCD and conditional or normal TCD values. More recently, Berg (2012) published findings which suggest that children with SCD scored significantly lower than matched controls on the Metacognitive Index and Global

Executive Composite scales of the BRIEF. Due to the low number of studies revealing findings associated with neurologic test sensitivity and SCD morbidity, more research is needed to understand the utility of the BRIEF in assessing pediatric SCD morbidity.

Six research studies indicated that the Wechsler Intelligence Scales, Woodcock Johnson Cognitive Abilities, and D–KEFS were sensitive to the effects of pediatric SCD severity. This was demonstrated by the measures showing decreased performance associated with disease severity in comparison to healthy peers (Bernaudin, Verlhac, Freard, Roudot-Thoraval, Benkerrou, et al., 2000, Schatz, 2004; Schatz et al., 2004; Schatz et al, 2007; Kral, Brown, Nietert, Abboud, Jackson, MD*; and Hynd, 2003; Berg, Edwards, and King, 2012). Among this small group of measures the WJ-III were shown to have the strongest psychometric properties related to pediatric SCD, as reviewed above (see Table 1.4).

It was also shown that the D–KEFS and WJ-III indicated decreased cognitive performance associated with SCD status alone in five research studies (Goonan et al., 1994; Puffer, Schatz, Roberts, 2007; 2010; Schatz et al., 2005). These findings are believed to reflect disease-related neurologic problems that do not reveal themselves with current clinical neuroimaging methods. As the D–KEFS is the only measure within this body of literature to reflect the fairness guidelines for test practices, it is notable that results converge for both the WJ-III and this measure. Additionally, the Wechsler Scales, WJ-III, CVLT–II, and Wisconsin Card Sorting Test demonstrated test sensitivity for children with SCD and evidence of CVA within 13 research studies. Children with SCD and CVA evidenced greater deficits than healthy peers and children with SCD and silent infarct on these measures. However, as demonstrated by table 1.4 and 1.5, the Wisconsin

Card Sorting Task psychometrics was found to be low as well as the norm sample "poor." Consequently test use for SCD and CVA is not recommended for this measure.

Approximately a third of the research articles (n = 10) provided results which indicate children with SCD and silent infarct demonstrated cognitive performance which was sensitive to detection on the Wechsler Scales, WJ-III, CVLT–II, and the Children's Memory Scale. Children with SCD and silent infarct performed poorer than age-matched peers and better than children with SCD and stroke on these neuropsychological measures. Evaluation of the psychometrics and norm quality revealed that use of these tests is supported for all three of these measures however; the WJ-III might be preferred due to a better quality normative sample.

Additionally, across 19 studies, children with SCD and no abnormalities on MRI revealed cognitive dysfunction which could be demonstrated on the following scales: Wechsler Scales, WJ-III, Test, CVLT–II, Wisconsin Card Sorting Test, Trail Making Test, KABC, D–KEFS, and the Children's Memory Scales. With the exception of the Trail Making Test, each of these measures is supported for use with children diagnosed with SCD and cognitive deficits having normal MRI results. The exclusion of this measure is based on the psychometric properties for the Trail Making Test. They were revealed as adequate, but the normative sample was categorized as "poor" (see Table 1.4).

Neuropsychological measurement in pediatric SCD has limitations and strengths associated with the potential use of particular measures. Limitations such as those associated with utilizing tests such as Wisconsin Card Sorting and Trail Making Test with this population are based on limitations in the test development, such as norm

quality. The norm group for both measures is "poor," which suggests that comparisons made to these groups could result in inaccurate measurement of cognitive performance, particularly for clinical decision making at the individual level. Additionally, the psychometric properties reviewed only reached a threshold of adequate for these measures. Acceptable reliability and validity does not promote recommendation for use in the population when measures that achieved higher levels of reliability and validity are available.

Across studies which address morbidity associated with pediatric SCD the WJ-III and Wechsler Scales demonstrate the most consistent measurement sensitivity. However, the psychometric and normative sample quality of the WJ-III surpasses the Wechsler Scales for use with the pediatric SCD population (see Tables 4 and 5). Additionally, when selected for use, the D–KEFS was found to detect SCD related morbidity at the same rate as the WJ-III; however, the WJ-III is appropriate for a wider age range of children (preschool through older adults) whereas the Delis Kaplan Executive Function System is normed beginning at age eight years. The overall implication of this review of psychometrics is that among measures used to date, the WJ-III excels in all areas except in its use of a traditional, achievement approach to ability measurement.

Summary

The cognitive measurement practices for pediatric SCD studies and the examination of psychometric properties for these cognitive measures associated with pediatric SCD was identified. The purpose of these goals was to inform recommendations for neuropsychological assessment batteries to use in future SCD research and treatment studies. An examination of psychometric properties provided strong support for the

assertion that the WJ-III was the most appropriate of the current cognitive assessment batteries for use with pediatric SCD.

The most prominent theme demonstrated was the considerable inconsistencies in tests use for neuropsychological assessment of children with SCD. Use of varied cognitive measurement instruments has impacted the assessment of EF within both research and treatment methodological designs. Table 1.1 revealed that 57 different psychological measures have been used to better understand cognition in this special population. Due to the wide range of instruments in the pediatric SCD research literature, treatment research has been uninformed regarding the most appropriate cognitive construct to focus on and which instruments should be used to measure the construct. The use of many of these neuropsychological instruments in making clinical decisions would be substandard scientific practices. Review of any psychometric tool prior to administration is best practices for test administration (Groth-Marnat, 2003). It is imperative that future research related to this body of literature utilize standard instruments with more regularity to assess cognitive functioning the for pediatric SCD population. This uniformity will promote a better understanding of cognition in pediatric SCD.

It is also essential that treatment studies diverge from the common practice of utilizing global IQ scores to assess cognitive performance in pediatric SCD. Future SCD research efforts should take steps towards formulating a more comprehensive understanding of both the broad and narrow abilities associated with EF. This transition is important because children with SCD may demonstrate equivalent full scale IQ scores pre/post treatment yet still have deficits in narrow abilities (Schatz et al., 2002). Left un-

rehabilitated neurologic complications associated with pediatric SCD, have been shown to impact educational attainment (King et al., 2006). Additionally, use of measurement tools which assess a more complete picture of EF may further understanding of the cognitive profile for these children.

Among the identified measures within the pediatric SCD literature, three cognitive measurement batteries were utilized most often to evaluate neuropsychological ability: The Wechsler scales, WJ-III, and the Differential Ability Scales (DAS) assessment tools were shown to measure different components of cognition. These neuropsychological tools focus on different factors related to cognitive theory and diverse statistical practices to assess cognition. The WJ-III however, was the only theoretically driven assessment tool which was designed to measures discrete cognitive abilities. The other tools were largely developed to measure overall IQ with later delineation of narrow abilities that might be assessed with the subtests. Therefore, the theoretical orientation and capacity to measure various cognitive domains makes it a superior measurement battery in comparison to the Wechsler Scales and the DAS. Furthermore, the WJ-III is a good neuropsychological test to measure cognitive outcomes in treatment-focused pediatric SCD studies. The WJ-III recommendation is due to relatively easy administration procedures, better norms, and psychometrics for use with children who are ethnic minority or from lower income backgrounds (two demographic factors overrepresented in the SCD population). Thus, this cognitive battery should be utilized as the standard of comparison for neuropsychological measurement in the pediatric SCD field.

Future directions for neuropsychological measurement in pediatric SCD field include improvement in four areas. The first suggestion is that WJ-III should be administered as the best available measure to assess cognition in pediatric SCD research and treatment studies. Given the superior psychometric properties and excellent norm sample quality, this measure could further understanding of cognition in pediatric SCD. Secondly, measurement of narrow cognitive abilities should be included as outcomes for research and treatment studies instead of broad level IQ scores. This shift from global to specific measurement of cognition might help clinicians to glean a greater understanding of cognition for children with pediatric SCD and may better delineate treatment benefits. Finally, more sophisticated neuropsychological measurement tools should be developed to assess cognition for children with SCD. For example, a limitation of the WJ-III is that the underlying theory does not incorporate current knowledge of neuroscience and the constructs used may not map well onto populations with neurologic deficits and therefore may not be well suited to measuring specific changes in brain function (e.g., in treatment studies). A useful approach may be to compare the WJ-III to experimental measures of EF that have well known neural correlates. Comparative analysis of experimental measures to this "standard" will help the field develop more accurate measurement for children with pediatric SCD. The purpose of the present research study is to address afore mentioned suggestions.

An experimental EF assessment tool, the "EXAMINER" will be utilized to compare sensitivity and specificity relative to the WJ-III, measure cognitive narrow abilities, and asses its appropriateness for use with the pediatric SCD population. More specifically, this new EF battery will be psychometrically evaluated, assessment of its

differential sensitivity compared with the WJ-III, and the cultural validity of this instrument will be investigated for the measurement of cognitive deficits associated with pediatric SCD.

Table 1.1

Overview of Executive Function Measures Included in Sickle Cell Disease Literature

Reference	Neuropsychological Test	Acronyms	EF Domain
Thompson et al., 2002	Bayley Scales of Infant Development III	BSID-II	Mental, Motor, and Behavioral Development
Puffer et al., 2010 Brown et al., 2000 Cohen et al., 1994 Brown et al., 1993 Swift et al., 1989	Beery-Buktenica Developmental Test of Visual-Motor Integration	B-VMI	Visual-Motor Deficits
Schatz et al., 2004[i] Schatz et al., 2001 Schatz et al., 1999 DeBaun et al., 1998	Benton Judgment of Line Orientation	BJLO	Visuospatial Judgment
Craft et al., 1993	Benton Visual Retention Test	BVRT	Visual Perception, Visual Memory, Visuoconstructive Abilities
Craft et al., 1993	Benton Word Fluency Test	BWF	Speed and Ease of Verbal Production
Brown et al., 2000 Schatz et al., 2001 Cohen et al., 1994	Boston Naming Test	BNT	Word Retrieval Performance
Berg et al., 2012 Kral et al., 2003 Kral et al., 2004	Behavior Rating Inventory of Executive Function	BRIEF-Parent Form	Parent Observation of Executive Function Behaviors
White et al., 2006 Brandling-Bennett et al., 2003 Schatz et al., 2001 Schatz et al., 1999 DeBaun et al., 1998	California Verbal Learning Test – Children	CVLT-C	Verbal Learning and Memory
Schatz et al., 2001 Brown et al., 2000	Cancellation of A's Task	CAT	Visuospatial Function and Attention
Schatz et al., 1999 Craft et al., 1993	Children's Auditory Verbal Learning Test	CAVLT	Verbal Learning and Memory
Berg et al., 2012	Children's Kitchen Task Assessment	CKTA	Executive Functioning
Kral et al., 2003 White et al., 2006 Brandling-Bennett et al., 2003	Children's Memory Scale	CMS	Learning and Memory
Brown et al., 1993	Conner's Continuous Performance Task	CPT	Attention and Impulsivity

Reference	Neuropsychological Test	Acronyms	EF Domain
Kral et al., 2003	Conner's Continuous Performance Task II	CPT-II	Attention and Impulsivity
Craft et al., 1994	Covert Orienting Task	COT	Sustained Visual Task
Schatz et al., 2007	Delayed Response Task	DRT	Spatial Working Memory and Inhibition Processes
Berg et al., 2012 Puffer et al., 2007 Schatz et al., 2006 Schatz et al., 2005 Schatz et al., 2004[ii] Schatz, 2004	Delis-Kaplan Executive Function System	DKFS	Higher Level Cognitive Functions
Cohen et al., 1994 Swift et al., 1989	Detroit Test of Learning Aptitude II	DTLA-II	Specific Mental Abilities
Schatz et al., 2007	Denver Developmental Screening Test II	DDST-II	Cognitive and Behavioral Problems
Tarazi et al., 2007	Developmental NeuroPsychological Assessment	NEPSY	Neuropsychological Development
Steen et al., 2002	Developing Skills Checklist	DSC	Wide Range of Skills
Tarazi et al., 2007 Schatz et al., 2004[a] Schatz et al., 2002 Schatz et al., 2001 Schatz et al., 1999 DeBaun et al., 1998	Differential Ability Scales	DAS	Variety of Cognitive Abilities
Schatz et al., 2004[a]	Divided Hierarchical Attention Task	DHAT	Divided Attention
Brown et al., 1993	Expressive One Word Vocabulary Test	EVT	Oral Language
Cohen et al., 1994 Craft et al., 1993	Finger Tapping	FT	Motor Speed
Cohen et al., 1994	Finger Tip Number Writing	FTNW	Complex Tactile Perception and Concentration
Cohen et al., 1994	Gray Oral Reading Test – Revised	GORT=R	Oral Reading Growth and Difficulties
Brown et al., 1993 Fowler et al., 1988	Kagan Matching Familiar Figures Test	MFFT	Attention and Impulse Control (Cognitive Tempo)
Goonan et al., 1994	Gordon's Continuous Performance Task	GCPT	Selective Attention and Impulsivity
Puffer et al., 2010 Brown et al., 2000 Cohen et al., 1994 Brown et al., 1993 Swift et al., 1989	Kaufman Assessment Battery for Children 2nd ed.	KABC-II	Cognitive Development
Schatz et al., 2004 [a]	Koenig, Reiss, and Kosslyn's Categorical and Coordinate Spatial Judgment Task	KRKCSI	Categorical and Coordinate Spatial Judgment

Reference	Neuropsychological Test	Acronyms	EF Domain
Bernaudin et al., 2000	l'Alourette French Reading Test	LFMT	Reading Achievement
Wasserman et al., 1991	Luria-Nebraska Neuropsychological Battery	LNNB	Neuropsychological Impairment
Schatz et al., 2007	MacArthur Communication Development Inventory	CDI	Language and Communication Skills
White et al., 2000	Memory Span	MS	Short Term Memory
Craft et al., 1993	Mini Mental Status Examination	MMSE	Mental Impairment
Schatz et al., 2004 [a] Schatz et al., 2001 Brown et al., 1993 Schatz et al., 1999 DeBaun et al., 1998 Brown et al., 1993	Peabody Picture Vocabulary Test Revised	PPVT-R	Receptive Language
Tarazi et al., 2007 Bernaudin et al., 2000 Brown et al., 2000 Armstrong et al., 1996 Cohen et al., 1994	Purdue Pegboard Test	PPT	Fine and Gross Motor Dexterity and Coordination
Schatz et al., 2001 Brown et al., 2000	Rapid Automatized Naming	RAN	Ability to Read, Phonological Awareness, and Verbal IQ,
Schatz et al., 2006 Schatz et al., 2005	Self-Ordered Pointing Test	SOPT	Behavior by Using Plans and Strategies
Schatz et al., 2004 [a] DeBaun et al., 1998	Simple Reaction Time	SRT	Motor Skills
Brown et al., 1993	Stanford Binet IV	SB-IV	Cognitive Ability
Schatz et al., 2001 Schatz et al., 1999 DeBaun et al., 1998	Test of Variables of Attention	TOVA	Attention Disorders
Cohen et al., 1994	Test of Visual Perceptual Skills	TVPS	Visual Perceptual Strengths and Weaknesses
Puffer et al., 2010	Test of Language Development – Primary Version, 3rd ed.	TOLD III	Language Development
Cohen et al., 1994	Test of Visual-Perceptual Skills	TVPS	Visual Perceptual Strengths and Weaknesses
Schatz et al., 1999	Tower of Hanoi	TOH	Problem Solving, Executive Function Deficits
Kral et al., 2003	Trail Making Test	TMT	Visual Attention and Task Switching

Reference	Neuropsychological Test	Acronyms	EF Domain
Puffer et al., 2007 Schatz et al., 2006 Steen et al., 2005 Schatz et al., 2005 Schatz, 2004 Schatz et al., 2004[b] Steen et al., 2003 Boni et al., 2001 Schatz et al., 2001 Wang et al., 2001 Bernaudin et al., 2000 Brown et al., 2000 Steen et al., 1999 Steen et al., 1998 Watkins et al., 1998	Wechsler Intelligence Scale for Children 3rd ed.	WISC-III	Cognitive Ability
Berg et al., 2012	Wechsler Intelligence Scale for Children 4th ed.	WISC-IV	Cognitive Ability
Steen et al., 2003 Wang et al., 2001 Steen et al., 1998 Armstrong et al., 1996 Knight et al., 1995 Cohen et al., 1994 Craft et al., 1993 Wasserman et al., 1991 Swift et al., 1989	Wechsler Intelligence Scale for Children Revised	WISC-R	Cognitive Ability
Watkins et al., 1998	Wechsler Memory Scale	WMS	Memory Functions
Watkins et al., 1998	Wechsler Preschool and Primary Scale of Intelligence Revised	WPPSI-R	Cognitive Ability
Tarazi et al., 2007	Wechsler Preschool and Primary Scale of Intelligence 3rd ed.	WPPSI-III	Cognitive Ability
Kral et al., 2003 Kral et al.,2004 White et al., 2006 Schatz et al., 2002 Knight et al., 1995	Wechsler Abbreviated Scale of Intelligence	WAIS	Cognitive Ability
White et al., 2006	Wechsler Individual Achievement Test	WIAT	Academic Strengths and Weaknesses
Cohen et al., 1994	Wepman Auditory Discrimination Test	WADT	Auditory Discrimination
Schatz et al., 2004[b]	Woodcock-Johnson Psychoeducational Test Battery	WJ	Cognitive Abilities, Scholastic Aptitude, and Academic Achievement
Puffer et al., 2010	Woodcock-Johnson Psychoeducational Test Battery III	WJ-III	Cognitive Abilities, Scholastic Aptitude, and Academic Achievement

Reference	Neuropsychological Test	Acronyms	EF Domain
Kral et al., 2003 Puffer et al., 2007 Schatz et al., 2005 Schatz, 2004 Schatz et al., 2004 [a] Wang et al., 2001 Schatz et al., 2001 Brown et al., 2000 Schatz et al., 1999 DeBaun et al., 1998 Armstrong et al., 1996 Craft et al., 1993 Swift et al., 1989	Woodcock-Johnson Psychoeducational Test Battery Revised	WJR	Cognitive Abilities, Scholastic Aptitude, and Academic Achievement
Cohen et al., 1994 Wasserman et al., 1991 Fowler et al., 1988	Wide Range Achievement Test Revised	WRAT-R	Achievement Test
Schatz et al., 2001	Wide Range Assessment of Memory and Learning	WRAML	Memory Functions
Schatz et al., 2001 Schatz et al., 1999 DeBaun et al., 1998	Wisconsin Card Sorting Test	WCST	Ability to Display Flexibility in the Face of Changing Schedules of Reinforcement
Berg et al., 2012	Children's Kitchen Task Assessment	CKTA	Initiation, Sequencing, Safety Judgment, Organization, Working Memory

Cognitive tests utilized for analysis of Executive Functioning across pediatric SCD literature.

Table 1.2.

Descriptive Table of Pediatric Sickle Cell Disease Studies with Executive Function Measures

Reference	Measure	Norms: Participant Group Characteristics	Age SCD (mean)	*n* SCD	*n* Controls
Armstrong et al., 1996	Purdue Pegboard Test Wechsler Intelligence Scale for Children Revised Woodcock Johnson Revised	Std Norms: SCD – Normal MRI vs. Silent Infarct vs. CVA	6.1	*n* = 194	*n/a*
Berg et. al., 2012	Delis-Kaplan Executive Function System Wechsler Intelligence Scale for Children IV Behavior Rating Inventory of Executive Function Children's Kitchen Task Assessment 144: French Memory Scale	Local/Std Norms: SCD – Normal	10.2	*n* = 22	*n* = 22
Bernaudin et al., 2000	Batterie d'Efficience Mnesique l'Alourette French Reading Test Purdue Pegboard Test Wechsler Intelligence Scale for Children III	Local/Std Norms: SCD – Abnormal MRI vs. Abnormal TCD vs. Low Hematocrit vs. Siblings	10.2	*n* = 173	*n* = 76
Boni et al., 2001	Wechsler Intelligence Scale for Children III	Std Norms: SCD – CVA and HbSS vs. Normal MRI and HbSS vs. Normal MRI and HbSC	11.4	*n* = 52	n/a
Brandling-Bennett et al., 2003	California Verbal Learning Test for Children – Children's Version Children's Memory Scale	Std Norms: SCD - Abnormal MRI vs. Normal MRI	12.9	*n* = 83	n/a
Brown et al., 2000	Cancellation A's Task Trails Making Test (B) Wechsler Intelligence Scale for Children III Woodcock Johnson Revised	Local/ Std Norms: SCD - CVA vs. Silent stroke vs. Normal MRI	9.8	*n* = 63	n/a

Reference	Measure	Norms: Participant Group Characteristics	Age SCD (mean)	*n* SCD	*n* Controls
Brown et al., 1993	Beery-Buktenica Developmental Test of Visual-Motor Integration Continuous Performance Task Expressive One Word Vocabulary Test Kagan Matching Familiar Figures Test Kaufman Assessment Battery for Children 2nd ed. Peabody Picture Vocabulary Test Revised Picture Vocabulary Test Stanford Binet IV	Local/ Std Norms: SCD vs. Siblings	10.4	*n* = 88	n/a
Cohen et al., 1994	Beery-Buktenica Developmental Test of Visual-Motor Integration Boston Naming Test Finger Tapping Finger Tip Number Writing Kaufman Assessment Battery for Children 2nd ed Purdue Pegboard Test Test of Visual-Perceptual Skills Wechsler Intelligence Scale for Children Revised Wepman Auditory Discrimination Test Wide Range Achievement Test Revised	Local/ Std Norms: SCD - Left Cerebral Infarct vs. Right Cerebral Infarct	9.96 years	*n* = 10	n/a
Craft et al., 1994	Covert Orienting Task	Local: SCD - no CVA vs. Bifrontal Lesions vs. Diffuse Lesions vs. Siblings	10.7	*n* = 29	*n* = 20
Craft et al., 1993	Benton Verbal Fluency Test Benton Visual Retention Test Children's Auditory Verbal Learning Test Finger Tapping Mini Mental Status Examination Wechsler Intelligence Scale for Children Revised Woodcock Johnson Revised	Local/Std Norms: SCD – CVA with Anterior Lesions vs. CVA with Diffuse Lesions vs. CVA with Normal MRI vs. Siblings	7.51	*n* = 39	*n* = 20
DeBaun et al., 1998	Benton Judgment of Line Orientation Benton Visual Form California Verbal Learning Test – Children Choice Reaction Time Task Differential Ability Scales Multilingual Aphasia Examination Peabody Picture Vocabulary Test Revised Simple Reaction Time Test of Variables of Attention Wisconsin Card Sorting Test Woodcock Johnson Revised Discrimination	Local/Std Norms: SCD – Silent Stroke vs. Overt Stroke vs. Siblings	11.5	*n* = 28	*n* = 17

Reference	Measure	Norms: Participant Group Characteristics	Age SCD (mean)	*n* SCD	*n* Controls
Fowler et al., 1988	Beery-Buktenica Developmental Test of Visual-Motor Integration Kagan Matching Familiar Figures Test Wechsler Intelligence Scale for Children Revised Wide Range Achievement Test	Std Norms: SCD – No Stroke vs. Healthy Peers	12.8	*n* = 28	*n* = 28
Goonan et al., 1994	Gordon's Continuous Performance Task Kagan Matching Familiar Figures Test	Local/ Std Norms: SCD – No CVA vs. Sibling	11.5	*n* = 24	*n* = 11
Knight et al., 1995	Wechsler Intelligence Scale for Children Revised Wechsler Abbreviated Scale of Intelligence	Std Norms: SCD vs. Healthy Peers	16.85	*n* = 60	*n* = 60
Kral et al., 2003	Behavior Rating Inventory of Executive Function Children's Memory Scale Conners's Continuous Performance Test II Trail Making Test Wechsler Abbreviated Scale of Intelligence	Std Norms: SCD vs. Normal	10.6	*n = 25*	*n = 35*
Kral et al., 2004	Wechsler Intelligence Scale for Children III Woodcock-Johnson Psychoeducational Test Battery Revised Behavior Rating Inventory of Executive Function: Parent/Teacher Wechsler Abbreviated Scale of Intelligence	Std Norms: SCD – TCD status Normal vs. Conditional vs. Abnormal	10.1	*n* = 62	*n* = 0
Puffer et al., 2010	Beery-Buktenica Developmental Test of Visual-Motor Integration Kaufman Assessment Battery for Children 2nd ed. Test of Language Development, 3rd. ed. Woodcock-Johnson Psycho-educational Test Battery III	Std Norms: SCD – SCD vs. Healthy Peers	6.5	*n* = 64	*n* = 81
Puffer et al., 2007	Delis-Kaplan Executive Function System Wechsler Intelligence Scale for Children III Woodcock Johnson Revised	Std Norms: SCD - Hydroxyurea Therapy vs. Without	13.2	*n* = 65	n/a
Schatz, 2004	Delis-Kaplan Executive Function System Woodcock Johnson Revised	Std Norms: SCD vs. Healthy Peers	11.6	*n* = 50	*n* = 36
Schatz et al., 2006	Delis-Kaplan Executive Function System Self-Ordered Pointing Test Wechsler Intelligence Scale for Children III	Local/ Std Norms: SCD - No Lesion vs. Silent Infarct vs. Overt Stroke vs. Healthy Peers	12.8	*n* = 28	*n* = 16

Reference	Measure	Norms: Participant Group Characteristics	Age SCD (mean)	*n* SCD	*n* Controls
Schatz et al., 2001	Benton Judgment of Line Orientation Boston Naming Test California Verbal Learning Test for Children – Children's Version Test of Variables of Attention Wisconsin Card Sorting Test Cancellation A's Task Differential Ability Scales Peabody Picture Vocabulary Test Revised Rapid Automatized Naming Trail Making Test Wechsler Intelligence Scale for Children III Woodcock-Johnson Psychoeducational Test Battery Revised	Local/Std Norms: SCD – Silent Cerebral Infarcts vs. Without Infarct vs. Siblings	10.8	*n* = 64	*n* = 16
Schatz et al., 2004[iii]	Benton Judgment of Line Orientation Differential Ability Scales Simple Reaction Time Woodcock-Johnson Psychoeducational Test Battery Revised Peabody Picture Vocabulary Test Revised Treisman and Gelade (1980) Parallel and Serial Visual Search Task Koenig, Reiss, and Kosslyn's Categorical and Coordinate Spatial Judgment Task	Local/ Std Norms: SCD - Unilateral Left Brain Injury vs. Unilateral Right Brain Injury vs. Bilateral Brain Injury vs. Siblings	12.8	*n* = 42	*n* = 43
Schatz et al., 1999	Benton Judgment of Line Orientation California Verbal Learning Test – Children Differential Ability Scales Peabody Picture Vocabulary Test Revised Test of Variables of Attention Tower of Hanoi Wisconsin Card Sorting Test Woodcock-Johnson Psychoeducational Test Battery Revised	Local/Std Norms: SCD – Anterior Cerebral Infarct vs. Diffuse Cerebral Infarct vs. Siblings	12.7	*n* = 28	*n* = 17
Schatz et al., 2004[iv]	Delis-Kaplan Executive Function System Wechsler Intelligence Scale for Children III Woodcock-Johnson Psychoeducational Test Battery Revised	Std Norms: SCD vs. Healthy Peers	11.6	*n* = 50	*n* = 36
Schatz et al., 2005	Delis-Kaplan Executive Function System Self-Ordered Pointing Test Wechsler Intelligence Scale for Children III Woodcock-Johnson Psychoeducational Test Battery Revised	Local/ Std Norms: SCD vs. Healthy Peers	11.4	*n* = 25	*n* = 25

Reference	Measure	Norms: Participant Group Characteristics	Age SCD (mean)	*n* SCD	*n* Controls
Schatz et al., 2007	Delayed Response Task Denver Developmental Screening Test II MacArthur Communication Development Inventory	Local/ Std Norms: SCD - High Neurologic Risk vs. Low Neurologic Risk	2.2	*n* = 61	n/a
Schatz et al., 2002	Differential Ability Scales Wechsler Abbreviated Scale of Intelligence	Std Norms: SCD - Silent Infarct Cases vs. No Infarct Cases	12	*n* = 25	n/a
Steen et al., 2005	Wechsler Intelligence Scale for Children Revised Wechsler Intelligence Scale for Children III	Std Norms: SCD vs. Healthy Peers	10.9	*n* = 54	*n* = 2200
Steen et al., 2002	Developing Skills Checklist	Std Norms: SCD vs. Healthy Peers	5.6	*n* = 34	*n* = 68
Steen et al., 2003	Wechsler Intelligence Scale for Children III Wechsler Intelligence Scale for Children Revised	Std Norms: SCD - Abnormal MRI vs. Normal MRI	9.5	*n* = 49	n/a
Steen et al., 1998	Wechsler Intelligence Scale for Children III Wechsler Intelligence Scale for Children Revised	Std Norms: SCD – No MRI Abnormalities vs. Abnormal MRI vs. Healthy Peers	8.4	*n* = 30	*n* = 24
Steen et al., 1999	Wechsler Intelligence Scale for Children III	Std Norms: SCD vs. Healthy Peers	10.6	*n* = 50	*n* = 52
Swift et al., 1989	Wechsler Intelligence Scale for Children Revised Woodcock-Johnson Psychoeducational Test Battery Revised Kaufman Assessment Battery for Children 2nd ed. Detroit Test of Learning Aptitude II Beery-Buktenica Developmental Test of Visual-Motor Integration	Std Norms: SCD with no CVA vs. Siblings	11.5	*n* = 21	*n* = 21
Tarazi et al., 2007	Differential Ability Scales Wechsler Preschool and Primary Scale of Intelligence Revised PPT Developmental NeuroPsychological Assessment	Std Norms: SCD – No History of Stroke	4.3	*n* = 28	n/a
Thompson et al., 2002	Bayley Scales of Infant Development III	Std Norms: SCD – HbSS vs. HbSC/Other	1.8	*n* = 89	n/a

Reference	Measure	Norms: Participant Group Characteristics	Age SCD (mean)	*n* SCD	*n* Controls
Wang et al., 2001	Wechsler Intelligence Scale for Children III Wechsler Intelligence Scale for Children Revised Woodcock-Johnson Psychoeducational Test Battery Revised	Std Norms: SCD - Abnormal MRI vs. Normal MRI	8.1	*n* = 373	n/a
Wasserman et al., 1991	Luria-Nebraska Neuropsychological Battery Wechsler Intelligence Scale for Children Revised Wide Range Achievement Test Revised	Std Norms: SCD vs. Siblings	11.5	*n* = 43	*n* = 30
Watkins et al., 1998	Wechsler Intelligence Scale for Children III Wechsler Memory Scale Wechsler Preschool and Primary Scale of Intelligence Revised Wisconsin Card Sorting Test	Std Norms: SCD – CVA vs. No CVA and Silent Infarct vs. No CVA and Normal MRI vs. Siblings	10.62	*n* = 41	*n* = 15
White et al., 2006	Wechsler Abbreviated Scale of Intelligence Wechsler Individual Achievement Test California Verbal Learning Test – Children Children's Memory Scale	Std Norms: SCD - Abnormal MRI vs. Normal MRI	11.95	*n* = 65	n/a
White et al., 2000	Memory Span	Local: SCD: History of Stroke vs. No History)	11.7	*n* = 31	n/a

Descriptive information for the selected studies utilized in the literature review.

Table 1.3.

Review of the Prominent Executive Function Psychometrics

Measure & Subtest	Type of Test	Age Range	Test-Retest Reliability	Internal Consistency	Convergent Validity	Norm Quality	Reference
Behavior Rating Inventory of Executive Function	Traditional	5-18 years	Good	Good	Child Behavior Checklist	Adequate	Gioia, et al. 2000
California Verbal Learning Test – Children	Traditional	5-16 years	Marginal		California Verbal Learning Test: Moderate	Adequate	Diller, 1974
-Long Delay Free Recall				Excellent			Van der Elst et al., 2005
Cancellation A's Task	Subtractive Factors			DATA NOT REPORTED			
Children's Auditory Verbal Learning Test	Traditional	7-17 years	Good	Good	CVLT: Strong	Good	Kennepohl et al., 2004 Schmidt, 1996 Tally, 1990
Children's Memory Scale	Traditional	5-16 years			Wechsler Memory Scale III: Strong	Adequate	Cash, 2009 Cohen, 1997 Strauss et al., 2006
-Backwards Digit Span			High	High			

Measure & Subtest	Type of Test	Age Range	Test-Retest Reliability	Internal Consistency	Convergent Validity	Norm Quality	Reference
Children's Kitchen Task Assessment	Traditional	8-12 years	Moderate	High	BRIEF: Parent	Poor	Rocke et al., 2008
Choice Reaction Time Task	Subtractive Factors	DATA NOT REPORTED					Canavan et al., 1974
Conner's Continuous Performance Task	Subtractive Factors	6-55+ years	Adequate	Not Reported	Posner Visual Orienting Task: Strong	Poor	Conners, 1994
-II	Subtractive Factors	6-55+ years	Moderate	Modest	Beery Test of Visual Perception: Acceptable	Poor	Strauss, Sherman, & Spreen, 2006
Memory Span	Traditional	DATA NOT REPORTED					Gihooly et al., 1980
Mental Status Examination – Auditory Vigilance	Interview	DATA NOT REPORTED					Strub et al., 1977
Multilingual Aphasia Examination	Traditional	6-12 and 16-97 years	DATA NOT REPORTED		Boston Naming Test and Wechsler Abbreviated Scale of Intelligence Revised – Verbal subtests: High	Poor	Schum et al., 1989
Odd Man Out	Traditional	22-79	DATA NOT REPORTED		Cattell Culture Fair Intelligence Quotient: Strong	Poor	Pomati et al., 1996 Diascro et al., 1994 Frearson et al., 1986
Self-Ordered Pointing	Subtractive Factors	5-11 years	Adequate	High	Go-No-Go: Moderate	Poor	Strauss et al., 2006 Archibald et al., 1999 Petrides et al., 1982 Green et al. 1990
Stanford Binet IV	Traditional	2-23 years			Kaufman Assessment Battery for Children: Strong	Good	Thorndike et al., 1986

Measure & Subtest	Type of Test	Age Range	Test-Retest Reliability	Internal Consistency	Convergent Validity	Norm Quality	Reference
Test of Variable of Attention	Traditional	4-19 years		Moderate to High	Criteria for ADHD in the *Diagnostic and Statistical Manual of Mental Disorders* 4^{th}: Strong	Poor	Riccio et al., 2001 Dupuy Cenedela, 2000 McCarney et al., 1990
-Reaction Time			Very High				
Tower of Hanoi	Puzzle	7+	Low	Unacceptable	Tower of London: Low	Not Reported	Bishop et al., 2001 Welsh et al., 2000 Klahr et al., 1981
Trail Making Test	Subtractive Factors	9-14 and 15-89 years		Moderate to High	Digit Modality Test: Moderate	Poor	Strauss et al., 2006 Lucas et al., 2005 Kennepohl et al., 2004 Lee et al., 2000
-B			Adequate	Acceptable	Wisconsin Card Sorting Test and Visual Search and Attention and Test: Strong		
Wechsler Abbreviated Scale of Intelligence	Traditional	6-89 years	Good	Good	Kaufman Brief Intelligence Test: Strong	Adequate	Wechsler, 1999

Measure & Subtest	Type of Test	Age Range	Test-Retest Reliability	Internal Consistency	Convergent Validity	Norm Quality	Reference
Wechsler Intelligence Scale for Children 3rd ed.	Traditional	6-16 years			WPPSI-R, WISC-R, WAIS-R, Differential Ability Scales: Very Strong	Adequate	Canivez et al., 1999 Wechsler, 1997
-Digit Span Backwards			High	Acceptable			
-Coding			Adequate	Excellent			
-Picture Arrangement			Adequate	Questionable			
-Symbol Search			Adequate	Moderate			
-Freedom from Distractibility			Adequate	Adequate			
-Processing Speed			Adequate	Adequate			
	Traditional	6-16	Adequate	Good	WPPSI-III, WISC-III, WAIS-III,: Very Strong	Good	Wechsler, 2003
			Adequate	Good			
-4th edition			Good	Good			
	Traditional	3-80 years			Stanford-Binet: Moderate	Poor	Woodcock, et al., 1977
-Digit Span Backwards							
-Digit Span Forwards							

Measure & Subtest	Type of Test	Age Range	Test-Retest Reliability	Internal Consistency	Convergent Validity	Norm Quality	Reference
Woodcock Johnson Psycho-Educational Battery							
-Visual Matching			Adequate	Moderate	Wechsler Intelligence Scale for Children		
-Spatial Relations			High	Moderate	Revised: Strong		
-Category Fluency			Adequate	Moderate			
-Digit Span			Marginal	Moderate			
Woodcock Johnson Psycho-Educational Battery III	Traditional	2-90 years			Standford Binet IV & Kaufman Adolescent and Adult Intelligence Test: Strong	Excellent	Woodcock, et al., 2001
-Decision Speed			High	Excellent			
-Memory for Words			Adequate	Good			
Woodcock Johnson Psycho-Educational Battery Revised	Traditional	2-90 years			Assessment Battery and Wide Range Achievement Test: Strong		Woodcock, et al., 1989
-Visual Matching			Adequate	Excellent			
-Processing Speed			Adequate	Excellent			
-Spatial Relations			High	Good			
Wisconsin Card Sorting Test	Traditional	5-89 years	Low to Marginal	DATA NOT REPORTED	Trails Making Test B: Marginal	Poor	Boone, 1999 Ariffa et al., 1998 Boone et al., 1993 Axelrod et al., 1992

Notes: Test Retest Reliability Coefficients: Very High (≥.90), High (.80-.89), Adequate (.70-.79), Marginal (.60-.69), Low (.50-.59), Extremely Low (≤.49)
Internal Consistency: Excellent (α ≥ .9), Good (.9 > α ≥ .8), Acceptable (.8 > α ≥ .7), Questionable (.7 > α ≥ .6), Unacceptable (.5 > α)
Convergent Validity Correlations: Strong (.50 to 1.0), Moderate (.30 to .49), Low (.10 to .29)
Norm Quality: Sample descriptive: number, age, geographic location, gender, race/ethnicity, education (parental), inclusion criteria (special education, gifted): high, adequate, poor

Psychometric properties of neuropsychological measures included in literature review.

Table 1.4.

Complete Normative Sample Psychometric Considerations

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Behavior Rating Inventory of Executive Function Gioia, et al., 2000	*n = 1,419*	Yes	Age Bands	No	Males generally higher ratings for Inhibition	Maryland public and private schools 26.5% urban, 59% suburban, 14.5% rural	Unknown	Yes	No	Exclusion: Special Education or Psychotropic Medication
California Verbal Learning Test – Children Delis et al., 1994	*n* = 920	Yes	Random Age, Location, Parent Education, Gender, Race/ Ethnicity	No	Minimal, overall females outperform males	Public and Private Schools in Western, North, Central, Northeast, Southern US	Unknown	No	Yes	Exclusion: English Fluency
Cancellation A's Task Diller, 1974	DATA NOT REPORTED									

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Children's Auditory Verbal Learning Test Van der Elst et al., 2005	Meta-Norms: Children *n* = 80-900	Yes	Age Bands	No	A female Advantage has been Shown in Recall and Recognition Trials	International	Lower Levels of Acculturation are Linked to Lower Scores	Yes	No	Exclusion: Learning Disabilities, Attention Disorders, Special Academic Assistance, Very High or Low Academic Performance
Children's Memory Scale Strauss et al., 2006	*n* = 1000	Yes	Age Bands	No	Not Reported	National	Minimal if Parent Education is Controlled	No	No	Exclusion: Below grade level reading, grade repetition, special education, neurologic disorder, and head injury
Choice Reaction Time Task Canavan et al., 1974	DATA NOT REPORTED									

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Conner's Continuous Performance Task Conners , 1994	*n* = 1190 Note: Clinical - *n* = 670	No	8 Age Bands	No	Not Reported	5 states and Southern Ontario	Not Reported	No	No	Exclusion: "Outliers" on Attentional Diagnoses and those Taking Psychotropic Medication
Controlled Oral Word Association Test Anderson et al., 1997	*n* = 422	No	Age Bands Note: Age Bands 14 and 15 contain less than 20 children	No	Overall Females Shown to Out-perform Males	Melbourne Australia	Scores shown to be less than Caucasian Peers	No	No	Exclusion: English as Second Language, History of Neurological, Sensory, Developmental Disorders
Delayed Response	DATA NOT REPORTED									
Delis-Kaplan Executive Function System Delis et al., 2001	*n* = 750	Yes	Random Age, Gender, Race/ Ethnicity, Education Level	No	Not Reported	38 States	Not Reported	No	No	Sensory Limitations, Substance Abuse, Medical, Psychiatric, or Motor Condition

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Detroit Test of Learning Aptitude 2nd Edition Hammill et al., 1985	*n* = 1532	No	Missing Information for Sample Ethnicity, Age, Education of parents	No	No	30 states	Unknown:Black and White Ethnicity is Reported Together as "Other" (92% of sample)	Not Reported	No	Not Reported
Divided Hierarchical Attention Task Robertson et al., 1993	DATA NOT REPORTED									
Gordon's Continuous Performance Task Gordon et al., 1988	*n* = 1300	Not Reported	Age Bands	Not Reported	Not Found	Syracuse, New York Charlottesville Virginia	Not Reported	No	No	Exclusion: ADD or ADD/H, Special Classroom Placement, Learning disability, IQ < 70, Major Neurological Impairment, Psychotropic Medication, Psychotherapy, Grade Retention, Emotional Disturbance

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Kaufman Assessment Battery for Children, 2nd ed. Kaufman et al., 2004	*n* = 3025	Yes	Random Age, Gender, Ethnicity, Parental Education, Geographic Region	Yes	Not Found	39 States and DC	Scores shown to be less than Caucasian Peers	No	Yes	Inclusion: Learning Disability, Speech Language Impairment, Mental Retardation, Emotional/Behavioral Disturbance, ADHD, Gifted
Kagan Matching Familiar Figures Test Salkind, et al., 1978	*n* = 5000 Note: 2,676 American	No	Age Bands	No	Female Scores Correlated with Verbal Intelligence	International	Not Reported	No	No	Inclusion: Normal Intelligence, Middle-Class SES
Memory Span Gilhooly et al., 1980	DATA NOT REPORTED									
Mental Status Examination Strub et al., 1977	DATA NOT REPORTED									

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Multilingual Aphasia Examination Schum et al., 1989	*n* = 229	Yes	Age Bands	No	Not Found	Iowa	Black, Urban, Inner City Culture and Lower Levels Acculturation have been Linked to Lower Scores	No	No	Inclusion: Children with Scores of 80-120 on the Peabody Picture Vocabulary Test -Revised included
Self-Ordered Pointing Test Petrides et al., 1982	*n* = 89	Yes	Age Bands	No	Not Found	Greater Victoria School District in Canada	Not Reported Note: Race/ Ethnicity Not Reported	No	No	Exclusion: Significant Neurological, Psychiatric, Developmental Difficulties
Stanford Binet IV Thorndike et al., 1986 Greene et al., 1990	*n* = 5013	Yes	Random Geographic Region, Community Size, Ethnic Group, Age, Gender, Socio-economic Status	No	Not Reported	National	Found to be Valid for Exceptional Black Male Students but not Learning Disabled	No	Yes	Exclusion: Severe Medical Conditions, Communication Deficits, Severe Behavior/ Emotional Disturbances

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Test of Variables of Attention Riccio et al., 2001 Dupuy Cenedela , 2000	*n* = 2551	No	Age and Sex	No	Males have been evidenced to have Faster Reaction Times and Errors in Scores for Subtests	Minneapolis Minnesota	Not Reported Note: Race/ Ethnicity of Sample: 99% Caucasian, 1% Other	No	No	Exclusion: Significant Neurological, Psychiatric, Developmental, or Learning Difficulties
Tower of Hanoi Welsh, 2000 Anderson, et al., 1996 Klahr et al., 1981	*n* = 200-300 Note: Individual Research Studies	Varied	Age Bands	No	Males Shown to Complete Task Quicker than Females	Varied	Not Reported	No	No	Exclusion: Varied
Trail Making Test Strauss et al., 2006 Lucas et al., 2005 Kennepohl et al., 2004	*n* = 100-392 Note: Meta-norms	Yes	Age Bands	No	Females Shown to have Minimal Advantage	International	Lower Overall Scores are Associated with Black English Speakers	No	Yes	Exclusion: Neurological, Developmental, or Psychiatric Disorders

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Wechsler Abbreviated Intelligence Scale ADD INFO										
Wechsler Intelligence Scale for Children 3rd Edition Canivez et al., 1999 Wechsler , 1997	*n* = 2200	Yes	Random Gender, Age, Parent Education	No	Female scores higher for Full Scale, Verbal, Performance FSIQ, Processing Speed	31 states	Black Scores shown to be less than Caucasian Peers	No	No	Exclusion: (Special Group) Gifted, and Mental Retardation, Reading
Wechsler Intelligence Scale for Children 4rd Edition Kaufman, et al., (2006)	*n* = 2,200	Yes	Random, Stratified	Yes	Females scored higher on Coding Males Scored higher on Information and Arithmetic No significant differences in FSIQ	4 geographical areas covering United States and Hawaii	Black Scores shown to be less than Caucasian Peers	No	No	Disability, Reading and Written Expression Disorder, Learning Disorder, ADD, ADHD, Expressive Language Disorder, Head Injury, Autistic Disorder, Motor Impairment:

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Wechsler Memory Scale Wechsler, 1945	*n* = 200 (adults)	No	Age Bands	No	Not Found	Bellevue Hospital, NY	Not Reported	No	No	Inclusion: "Normal" Adults
Woodcock Johnson Psycho-Educational Battery Woodcock et al., 1977	*n* = 3935	Yes	Age Bands	No	Females Out Preformed Males on Subtests Related to Verbal Performance	National Note: Rural N. Easterners and Southerners Under-represented	Scores shown to be less than Caucasian Peers	No	No	Exclusion: Not Reported
Woodcock Johnson Psycho-Educational Battery III Woodcock et al., 2001	*n* = 5972	Yes	Random Community, and Individual Variables Note: Variables -Census Region, Community size, Sex, Race, Hispanic, Type of school, Adult Education, Occupation & Status	No	Not Reported	100 Geographic Regions in 27 States	Scores have been Shown to have Comparable meaning for African American and Caucasian groups	No	No	Exclusion: Less than One Year Experience English Speaking Classes

Measure & Reference	Pediatric Sample Size	Sample Representative of US Census	Sample Stratified	Ethnicity Stratified by SES	Gender Effects	Sample Geographic Diversity	African American Cultural Effects	Sample General Health Condition Reported	Children Receiving Special Education Included	Sample Criteria
Woodcock Johnson Psycho Educational Battery Revised Woodcock et al., 1989	*n* = 4359	Yes	Age, Gender, Region, Community Size, and Race	No	Not Reported	Communities Selected by Socio-economic Variables in US	Not Found	No	Yes	Inclusion: Mainstreamed, Special Education, Handicapped
Wisconsin Card Sorting Test Heaton et al., 1993	*n* = 459	No	Age Bands	No	Not Reported Note: Larger Number of Males (58%)	National	Not Reported	No	No	Exclusion: Neurological Dysfunction, Learning Disability, Emotional and, Attention Disorders

Criteria used to evaluate the quality of the normative sample for executive function measure used to measure cognition in Pediatric SCD literature.

Table 1.5.

Studies by Type of SCD Morbidity

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Berg et al.,2012			X	X	X		Children with SCD earned lower scores in the domains of metacognition and global executive function on the BRIEF compared to healthy demographically matched controls. Children with SCD performed w, initiation, organization, and completion on the CKTA. Children with SCD revealed significant time differences for the Color-Word Interference test for all conditions except on DKFS except Naming Colors and the Sorting test.
Bernaudin et al., 2000	X	X		X	X	X	Children with SCD performed more poorly than siblings overall on WISC III Verbal IQ (Similarities, Comprehension, Digit Span), Performance IQ (Coding, Symbols), Processing Speed, and Full IQ. Children with SCD and overt stroke evidenced significant impairment on WISC III Performance IQ and Full Scale IQ compared to SCD with no history of stroke. Children with SCD and silent stroke performed poorer than children with no history of silent stroke for WISC III Similarities, Vocabulary, and Verbal Comprehension. Children with severely low hematocrit and those with thrombocytosis infarcts on magnetic resonance imaging evidenced poor scores on WISC III Verbal IQ, Performance IQ, and Full Scale IQ

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Brandling-Bennett et al., 2003		X		X			Children with SCD and frontal infarcts performed more poorly on learning and free recall tasks of the CVLT-C. Children with SCD and frontal infarcts revealed lower performance on Children's Memory Scale Backward Digit Span.
Brown et al., 2000		X		X	X		Children with SCD and no central nervous system pathology evidenced greater general performance than the SCD group with Silent Infarcts. Children with SCD and no history of central nervous system pathology evidenced less omission errors and data suggest there is a commission error trend on the Cancellations of As task. Children with SCD and localized frontal lobe lesions made significantly more omission and commission errors on the Cancellations of A's task.
Brown et al., 1993		X					Children with SCD evidenced greater errors on CPT than siblings. Children with SCD performed more poorly on K-ABC Reading Decoding and Bead Memory subtests than siblings. Older children with SCD performed worse B-VMI and MMFT than younger children with SCD.
Cohen et al., 1994					X		Children with SCD and left cerebral infarcts showed greater impairment on the Full Scale IQ, VIQ, and PIQ of the WISC-R than children with SCD and right cerebral infarcts. Children with SCD and left cerebral infarct showed global impairment on WRAT and GORT, Wepman, and KABC Number Recall when compared to children with SCD and right cerebral infarcts. Children with SCD and left cerebral infarct demonstrated impairment on Picture Naming for the Boston Naming Test.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Craft et al., 1994		X		X	X		Children with SCD and bi-frontal lesions evidenced quick disengagement at brief cue intervals and slower disengagement for lower intervals on the Covert Orienting Task. Children with diffuse lesions showed no lateralized effects and greater reaction time at brief cue delays on the Covert Orienting Task.
Craft et al., 1993		X		X	X		Children with SCD and diffuse cortical strokes performed more poorly on WISC-R Block Design, Benton Visual Retention, and WJR Spatial Relations than both children with SCD and anterior stroke group and siblings. Children with SCD and anterior lesions had greater intrusions during list recall on the CAVLT than both children with SCD and diffuse stroke and siblings. There were no significant differences between children with diffuse or anterior lesions and siblings on Finger Tapping, WISC- R Vocabulary, and CAVLT.
DeBaun et al., 1998				X	X		TOVA, Wisconsin Card Sorting Test, and the CVLT identified children with SCD and Silent Cerebral Infarct. TOVA (95%) evidenced greater sensitivity (86%) and specificity (81%) ratings than all other measures (≥ 60%). TOVA also demonstrated great sensitivity (**95%)** for children with SCD and overt stroke
Fowler et al., 1988		X					Children with SCD and no CVA evidenced no group differences for VIQ, PIQ, and Full Scale IQ from healthy peers on WISC-R. Children with SCD and no CVA demonstrated significant poorer performance than healthy peers on WISC-R Coding subtest. Children with SCD and no CVA performed more poorly on MFFT than healthy peers.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Goonan et al., 1994			X				Children's age contributed the greatest significance to predicted variance on CPT and MFFT performance. Younger children with SCD and no CVA showed poorer sustained attention on CPT and greater impulsivity on MFFT than older children with SCD and siblings.
Kral et. al., 2004						X	Children in the abnormal TCD group demonstrated greater deficits in the areas of inhibitory control, problem-solving flexibility, and modulation of emotional responses when compared to the conditional group, by parent observation (not significant). Teachers rated poorer performance for children with abnormal TCD findings in their ability to solve problems in working memory, plan and organize, as well as self-monitor.
Kral et. al., 2003						X	Children with abnormal TCD values evidenced lower performance than children with conditional TCD values in areas of verbal intelligence and executive function.
Puffer et al. 2010			X				Height for age in children with SCD was shown to partially account for poor cognitive performance on the TOLD III Spoken Language Quotient, B-DTVMI, KABC Hand Movements, WJ-III Decision Speed, Memory for Words, Letter Word Identification, and Applied Problems. Children with SCD and a greater body-mass-index performed significantly better on B-DTVMI and WJ-III Letter Word Identification and Applied Problems than children with SCD and smaller Body-mass-index. This result was not evidenced in healthy peers.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Schatz, 2004	X						Children with SCD had greater difficulties with WRAT-3 Single-Word Reading ability and Written Calculations than healthy peers. DKFS Digit Span–backward, WRAT-3 Single-Word Reading ability, and Written Calculations were valid predictors of functional education outcomes for children with SCD. Overall performance on cognitive measures was found to be a strong predictor for WRAT-3 Single-Word Reading ability and Written Calculations for children with SCD. Academic attainment was predicted by school absences related to SCD; however, academic achievement was not a significant predictor for illness related absences.
Schatz et al., 2001		X		X			Children with SCD and silent cerebral infarcts demonstrated a greater incidence of being retained a grade level in school or receiving special academic services, having cognitive deficits and frontal lobe injuries than children with SCD without infarcts. Language: Differential Abilities Scales Word Definitions, Word Fluency, WJ-R Picture Vocabulary, Peabody Picture Vocabulary Test, or Rapid Automatized Naming, and Boston Naming Test. Visual-spatial/visual-motor: Differential Abilities Scales Pattern Construction, Judgment of Line Orientation, and WJ-R Visual Closure, or Wechsler Block Design and Purdue Pegboard. Memory: CVLT-C or WRAML Attention and Executive Functioning: Test of Variables of Attention, Wisconsin Card Sorting Task, and Perseveration and Intrusion errors on the CVLT-C, or Cancellation of A's and the Trail Making Test.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Schatz et al., 2007	X						Older children with SCD showed poorer performance in language skills on the Denver-II, VABS Communication, CDI Vocabulary and on the fine motor scales of the Denver-II and VABS Motor. Working memory performance on the Delayed Response was poorer in children with higher risk SCD when compared to lower risk SCD.
Schatz et al., 2006					X		The corpus callosum of children with SCD and silent infarcts or overt stroke was smaller than children with SCD and no visible infarct or healthy peers. Lesion volume was a found to be a predictor of WISC-III Full Scale IQ, Verbal Comprehension, Perceptual Organization, Freedom from Distractibility, Processing Speed, DKFS Verbal Fluency, and Self-Ordered Pointing Test The corpus callosum rostral body size demonstrated robust prediction of performance on measures of WISC-III Processing Speed factor, DKFS Verbal Fluency, and the SOPT.
Schatz et al., 2005			X				Children with SCD demonstrated poorer performance for WJ-R Oral Vocabulary, Incomplete Words, Visual Matching and Digit Span-Backward on the WISC than healthy peers. Children with SCD demonstrated deficits in auditory processing, on the WJ-R Incomplete Words, suggesting relation to WISC Digit Span-Backward performance.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Schatz et al., 2004	X						Children with SCD and greater anemia severity performed poorer on DKEFS Category and Verbal Fluency, WISC III Full IQ, WJ-R Digit Span Backwards, Spatial Relations, and Visual Matching. Anemia severity was shown to predict overall cognitive ability, crystallized ability, and processing speed for children with SCD. A bio-psychosocial interaction of anemia severity and SES was found to predict general cognitive ability and short-term memory performance.
Schatz et al. 2004		X			X		Children with SCD and independent left or right cerebral injury demonstrate less efficient visual search for the contra-lateral visual field. Deficits in global-level processing measured by a task of divided hierarchical attention and coordinating spatial judgments assessed with Koenig, Reiss, and Kosslyn's Task were associated with right-hemisphere injury for children with SCD. Children with SCD and bilateral injury showed deficits in global-level processing and coordinate spatial judgments as well however they also evidenced a disruption of visual search across visual fields.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Schatz et al. 1999		X		X			Children with SCD and anterior cerebral infarcts, diffuse cerebral infarct, and siblings preformed significantly worse, respectively, on TOVA, Tower of Hanoi, Wisconsin Card Sorting Test, DAS Position, Visual Form, Shape Discrimination, Pattern Construction; Judgment of Line Orientation, and WJ-R Visual Closure. Children with SCD had generally lower scores on the Tower of Hanoi and TOVA compared to siblings. Volume of cerebral lesion was significantly related to performance on TOVA, Tower of Hanoi, WCST, CVLT-C, PPVT-R, and WJ-R Picture Vocabulary for children with SCD and anterior cerebral infarct.
Steen et al., 2005		X					Children with SCD and normal MRI showed significant deficits in WISC III Full-Scale IQ, VIQ, PIQ, Verbal Comprehension, Perceptual Organization, Freedom from Distraction, and Processing Speed compared with healthy peers. The performance of children with SCD and normal MRI decreased as a function of age for the WISC III Full Scale IQ, Verbal IQ, Verbal Comprehension, and Perceptual Organization.
Steen et al., 2003		X			X		Children with SCD and MRI abnormalities showed greater cognitive impairment than children with SCD and normal MRI in VIQ and Verbal Comprehension on WISC III. Children SCD and low hematocrit demonstrated greater impairment than children with SCD and higher hematocrit on Full Scale IQ, Verbal Comprehension, and Freedom from Distractibility.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Steen et al., 1999		X					Children with SCD and no CVA were found to be at a 23 fold increased risk of mild mental deficiency. Hematocrit of more than 27% in children with SCD was positively correlated with Full Scale IQ, Verbal Comprehension, Perceptual Organization, Freedom From Distractibility, and Processing Speed scores on WISC III. Hemtocrit of less than 27% was correlated with lower WISC III scores as well as T1 grey matter reductions in the caudate, putamen, and cortex regions of the brain.
Swift et al., 1989		X					Children with SCD and no CVA performed significantly poorer on DTLA-2 Memory for Words and Designs, WJ Math and Reading, and WISC-R VIQ, PIQ, Full Scale IQ. Children with SCD and no CVA along with healthy peers demonstrated that their WJ Math and Reading scores were commiserating with their WISC-R VIQ, PIQ, and Full Scale IQ performance.
Wang et al., 2001		X		X	X		Children with SCD and normal MRI had significantly higher scores on WISC III Full Scale IQ, VIQ, PIQ, Digit Span, Coding and WJ-R Math/Reading compared to children with SCD with Silent Infarcts or Stroke. Children with SCD and Silent Infarct performed significantly better than children with SCD and Stroke on WISC III Full Scale IQ, VIQ, PIQ, Digit Span, Coding and WJ-R Math/Reading. WISC III Verbal IQ, WJ-R Math Achievement, and WISC III Coding scores demonstrated decline with increasing age of the child with SCD.

Study Authors	SCD Morbidity						
	Disease Severity	Normal MRI	SCD Status	Silent Infarcts	Cerebral Vascular Accident	Abnormal TCD	Results
Watkins et al., 1998		X		X	X		Children with SCD and normal MRI demonstrated higher FSIQ, Digit Span, Coding, Picture Arrangement, and Symbol Search scores on WISC-III. A trend was evident for children with SCD and normal MRI suggesting greater performance on VIQ, PIQ, and FSIQ than SCD and CVA. Children with SCD and CVA demonstrated poorer performance overall on WISC-III compared to SCD and normal MRI and siblings. Children with SCD and CVA generated more perseverative errors on Wisconsin Card Sorting Test and had poorer performance on WMS Logical Memory, Visual Memory, and Paired Association Learning.
White et al., 2006		X		X			CVLT-C List A recall over 5 presentations was the most robust predictor for discriminating between children with SCD and silent infarct vs. no history of infarct. Block Design from the WAIS was the second best predictor of silent infarct in children with SCD. Together CVLT-C and Block Design from the WAIS demonstrated 75%, 75% sensitivity and 76% specificity for identifying children with SCD and silent infarct from children with SCD without silent infarct.

Schatz J, Craft S, Koby M, et al. (2004). Asymmetries in visual-spatial processing following childhood stroke. *Neuropsychology*, *18*, 340–352.

Schatz, J., Finke, R.L.,Roberts, C.W. (2004). Interactions among biomedical and environmental factors in cognitive development: A preliminary study of sickle cell disease. *Journal of Developmental and Behavioral Pediatrics*, *25*, 303–310.

Schatz J, Craft S, Koby M, et al. (2004). Asymmetries in visual-spatial processing following childhood stroke. *Neuropsychology*, *18*, 340–352.

Schatz, J., Finke, R.L., Kellett, J.M., & Kramer, J.H. (2002). Cognitive functioning in children with sickle cell disease: A metaanalysis. *Journal of Pediatric Psychology*, *8*, 739–748.

Schatz J, Craft S, Koby M, et al. (2004). Asymmetries in visual-spatial processing following childhood stroke. *Neuropsychology*, *18*, 340–352

[b] Schatz, J., Finke, R.L.,& Roberts, C.W. (2004). Interactions among biomedical and environmental factors in cognitive development: A preliminary study of sickle cell disease. *Journal of Developmental and Behavioral Pediatrics*, *25*, 303–310.

Note: Research studies which utilize cognitive measure that evidence measurement sensitivity to neurologic damage associate with pediatric SCD.

CHAPTER II

RESEARCH QUESTIONS AND HYPOTHESES

The development of the EXAMINER, a neuropsychological assessment battery that intended to supplant established cognitive measures such as the WJ-III and the BRIEF, is potentially important for more accurate EF assessment of children with SCD. Similar to the WJ-III, this newly developed measure includes subtests which evaluate the narrow domains of ability. However, these measures have been chosen to represent well established measures of EF, with known neural underpinnings and often used subtractive factors logic so that both overall performance and narrower abilities can be identified. Similar to the BRIEF, component tests can be combined to assess broader cognitive domains. As a new battery, it is not expected that the EXAMINER can match the WJ-III or the BRIEF in the area of a large, representative normative sample; however, if other features of the battery are shown to be promising, this expensive process could be implemented at a later point in time.

Three specific research questions will be addressed in this study: (1) Will the psychometric qualities of the EXAMINER meet or exceed the WJ-III standards based on: the structure of the test, target age range, reliability (internal consistency), size of normative sample, and quality of normative sample for children with SCD? (2) What is the extent of cross-cultural validity for the EXAMINER? (3) Does the EXAMINER demonstrate similar criterion validity as other well-established measures, such as the WJ-III, to detect Sickle Cell related cognitive deficits?

Study Hypotheses

Goal 1: Evaluate traditional psychometric properties of the EXAMINER relative to the WJ-III.

Hypothesis 1: The reliability (internal consistency) of the EXAMINER will be equivalent to the WJ-III for use in pediatric SCD, including the reliability for youth tested across multiple sites, local demographic controls, and youth with SCD.

The EXAMINER will be administered to a heterogeneous sample of children across multiple sites, a local demographic comparison group for our SCD sample, and a group of children with SCD in conjunction with selected measures from the WJ-III. The purpose of this comparison is to evaluate the reliability for the EXAMINER battery. It is hypothesized that the EXAMINER will be equivalent to the WJ-III on the demonstration of high internal consistency. The EXAMINER will also demonstrate better appropriateness for this population due to its inclusion of a special developmental sample of children with SCD (from this proposed study) supporting reliability within this clinical group.

Goal 2: Evaluate evidence of the cultural validity of the EXAMINER for African-Americans with SCD by comparing African-American youth relative to White youth on measures of convergent validity.

Hypothesis 2A: The strength of association between EXAMINER scores and VIQ from the WJ-III will be similar for African American and White children.

Hypothesis 2B: The strength of the association between the validity coefficient for the BRIEF Behavioral Regulation, Metacognition, and Global Executive Composite and

the EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor scores will be similar for African American and White children.

It is plausible that the EXAMINER will demonstrate similar or greater cultural validity for the WJ-III and BRIEF. The developers of the EXAMINER selected, developed, and piloted methods of assessment including neuropsychological tests, computer-administered tasks, and behavioral observation techniques with an *a priori* goal to develop measures that could be applied widely to individuals across ages and ethnicities. In addition, tasks were designed to minimize the need for specific cultural knowledge or specific linguistic knowledge that often is a source of bias in cognitive testing. Furthermore, the developers of the EXAMINER appropriately field tested and modified forms of tests and administration procedures for test takers. These modifications were made as developers received feedback from the national field testing sites that accommodations were necessary for standardization.

Goal 3: Establish the sensitivity (criterion validity) of the EXAMINER to SCD-related cognitive deficits.

Hypothesis 3A: Children with SCD at high risk for disease-related cognitive effects will show significantly poorer performance on the EXAMINER than demographically matched control children without SCD. The effect size of these differences will be compared to the magnitude of group differences on selected WJ-III.

Hypothesis 3B: The EXAMINER will show at least as high of sensitivity and specificity in identifying high risk children with sickle-cell disease as the selected measures from the WJ-III relative to the local, demographically matched comparison group.

It is believed that the degree of differences for the effect sizes of the EXAMINER, when compared with that found for established measures of cognitive functioning such as the WJ-III, will be similar or greater. With comparable criterion validity, the EXAMINER could be a preferred measure for clinical outcome studies (particularly treatment studies) given that the measures are based on well-established brain-behavior relationships, allows for both narrow and broad cognitive abilities to be measured, and specifically measures EF, a construct rooted in measuring higher-level cognitive skills that are essential for adaptive functioning.

CHAPTER III

METHODS

Participants

Clinical research participants were children who voluntarily took part in the EXAMINER national project (Kramer, Mungas, Possin, Rankin, Boxer, Rosen, Widmeyer, 2014). Overall, 396 children without a known neurologic condition (8 to 18 years of age) completed the study. One hundred and seventeen of these children, SCD and demographically matched controls, were recruited from a local catchment area surrounding Columbia, South Carolina. The overall sample for the multisite study contained 279 youth identifying as White who matched the age range of the local African-American children without known neurologic conditions. At the University of South Carolina site, 32 children diagnosed with SCD participated in the investigation. The 85 control participants were demographically matched on race and age for children for the SCD sample who voluntarily participated in the study from after school, summer, church, and community programs in the Columbia, SC area.

SCD subjects were a group of children diagnosed with high risk SCD (HbSS or HBSβ0). Children with known neurologic histories were included given the purpose of the study, though participants with severe sensory or motor deficits as judged by the treating hematologist were not recruited. Of the total SCD participants (n = 32), 56.3% were male (n = 18) and 43.8% were female (n = 14). The mean age of the SCD sample was 12.7 years (SD = 3.2). All local respondents self-identified as African American

ethnicity (*N* = 85), mean age 12.8 years (*SD* = 2.53). The local control sample was also relatively equally balanced for gender, 46.3% were male (*n* = 39) and 53.7% were female (*n* = 46). The multi-site sample (*N* = 279), contained participants which self-identified as White, mean age 11.2 years (*SD* = 2.9). Similar to the SCD and local control samples, the non-neurologic sample was comprised of 47.0% male (*n* = 133) and 53.0% female (*n* = 146). Fisher's Exact test indicated no differences across race/ethnicity groups for gender (p = .804). Socioeconomic (determined by parent age, education, income, number of children and adults in the home) information was also obtained for the University of South Carolina site participants. Unfortunately, this data was to have been collected from children at other sites, but was missing as described below.

General Inclusion/Exclusion Criteria

For the overall EXAMINER project there were a set inclusion/exclusion criteria across sites. Most of these criteria; however, were more relevant to adult neurologic conditions. General systemic medical conditions were exclusionary only if additional medical conditions were present besides the target condition, as with SCD. Participants had to be between 3-90 years old and speak fluent English and/or Spanish. Subjects who were unable to consent for themselves required an informant to consent for them. General exclusion criteria were current alcohol abuse or dependence, current drug abuse, psychiatric disorder (apart from those specified in diagnostic groups of interest), B12 deficiency or other metabolic syndrome, hypothyroidism (i.e. TSH>150% of normal), known HIV, renal failure, respiratory failure (i.e. requiring oxygen), significant systemic medical illnesses (e.g. deteriorating cardiovascular disease), or current

medication likely to affect CNS functions (e.g. benzodiazepines, antidepressants, lithium, and/or neuroleptics in the phenothiazine and haloperidol families).

Procedure

All procedures were approved by university and hospital institutional review boards and by the administrations of the schools and programs involved. Informed consent and assent were obtained from caregivers and children. The children diagnosed with SCD were recruited from Palmetto Health Children's Center for Cancer and Blood Disorders. Scheduling conflicts were most often cited as the reason for participation refusal. Children diagnosed with SCD with history of stroke or major developmental disabilities were excluded from analyses based on medical record review. Children experiencing any acute SCD complications did not participate in the investigation and were rescheduled at their next routine visit.

Neuropsychological assessment was conducted by a graduate student or trained undergraduate student under the psychologist's supervision. The data were collected via questionnaires with parents and one-on-one testing with the child. Parent questionnaires and BRIEF-Parent Form were collected and EXAMINERs were stationed nearby to address any parent questions or available to help the parent if questions arose. Children diagnosed with SCD were typically assessed after their routine care from the hematologist. Order of test administration was in a fixed order across the study. EXAMINERs were blind at the time of testing to disease-related measures (e.g., extent of SCD complications) and socioeconomic status, but not group diagnostic group (SCD vs. without SCD). Each child completed the EXAMINER and WJ-II administered in a single, 90- to 120-minute session. Compensation for participation involved the families

receiving monetary compensation (both parent and child) and SCD participants' data for clinically informative cognitive measures was shared with the hematologist.

The comparison group was recruited from the same general community to match the sample with SCD in terms of age, gender distribution, ethnicity, and SES. Caregivers were asked to enroll children in the study through telecommunications, electronic correspondence, and face to face contact at program sites. These sites consisted of local schools, after-school, summer, community, and church programs providing academic enrichment and leisure activities for children. Children with chronic health conditions or developmental disabilities were excluded from participation (based on parent report on a demographic survey). Demographically matched children also did not participate in the investigation during a time when they were experiencing acute illness.

Children without SCD were assessed at the end of their school day when the academic year was in session and various times of day during breaks from school. These participants were also assessed in a fixed order across the study. Parent report measures were completed by caregivers prior to study participation. Standardized procedures were utilized as described in the EXAMINER training manual.

Measures

Medical History

Medical chart reviews were conducted for all children with SCD to collect disease subtype. In addition, history of SCD complication (e.g. stroke, splenic sequestration) was reviewed to asses' inclusion/exclusion criteria for this study.

Socioeconomic Status

At our local site, demographic information was collected pertaining to age, gender, ethnicity, and the presence or absence of medical/psychiatric illness, or learning disability from participant's caregivers. Descriptive informative was also collected regarding age, gender, ethnicity, education, yearly household income, and household size from parents of participants. Household income data were collected in the following manner: less than $10,000 was recorded as Level One, $10,000 - $20,000 was recorded as Level Two, $20,000 - $30,000 was recorded as Level Three, $30,000 - $40,000 was recorded as Level Four, and more than $40,000 was recorded as Level Five. Parental education was recorded according to years of education and categorized: less than ninth grade recorded as Level One, ninth through eleventh grade recorded as Level Two, high school diploma/GED recorded as Level Three, Some college/Associates degree recorded as Level Four, and college degree/Bachelors or higher recorded as Level Five. Unfortunately, detailed demographic information was not collected at all sites in the overall study. It was intended for each of the 11 sites to collect parent education information for child participants. However, due to investigator error, education was only collected for adult participants. Therefore, we had access to information about age, gender, and race for children entered at other sites

Neuropsychological Assessment Tools

An experimental cognitive measurement test, the EXAMINER, WJ-III, and BRIEF-Parent Form were selected for use in this study to best capture EF in children with SCD.

Brief-Parent Form

The BRIEF-Parent Form is a parent questionnaire, created to measure the expression of the child's behaviors related to executive function. More specifically, it was included in this study to assess the observed cognitive abilities for a broad range of children ages 5–18 in the school and home environments. It was normed on the child ratings from 1,419 parents and has Eight theoretically derived clinical scales: Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. These scales form two broader Indexes: Behavioral Regulation and Metacognition were reviewed. An overall Global Executive Composite score, which depicts overall executive function, was also calculated from the clinical scales to provide a greater understanding of cognitive abilities for this population. Two validity scales: Inconsistency and Negativity were evaluated for each respondent. Gioia, Isquith, Guy, and Kenworthy (2000) reported high internal consistency (alphas = .80-.98) and test-retest reliability (*r*s = .82) for this measure.

Parents of participants were asked to complete an 86 items paper and pencil form. Each parent was asked to indicate if the behavior described in each item was never a problem, sometimes a problem, or often a problem for the child.

EXAMINER

The EXAMINER is a newly developed neuropsychological test battery that purports to reliably and validly assess domains of EF, across a wide range of disorders which impact cognition (Kramer et al., 2014). The test was designed largely using Miyake's unity and diversity model of EF with the expectation of capturing a global measure of executive function and more discrete component measures. A combination if

Item Response Theory (IRT) and factor analytic methods were used to extract useful factor scores from a set of well-established EF measures used in non-clinical contexts. This novel neuropsychological battery has components specifically designed to measure narrow domains of EF (e.g., attention, set shifting, inhibition, fluency, planning, problem solving, and abstract reasoning). Normative data has been derived from national standardization, involving children and adults between the ages of 5 and 89 years of age, including subjects with clinical syndromes known to impact cognition. Theoretically derived tasks were designed to measure narrow domains of EF; which include attention, set shifting, inhibition, fluency, planning, problem solving, and abstract reasoning. Performance on the EXAMINER tasks is used to inform fluency, cognitive control, and working memory factor scores, as well as an overall composite score. The fluency factor score is constructed from the total correct responses for 2 trials of Phonemic Fluency and 2 trials of Category Fluency tasks. The cognitive control factor score is comprised of the total Flanker score, total Set Shifting score, anti-saccade total, and total dysexecutive errors on the Continuous Performance Task. The Working Memory factor is comprised of total correct responses for the Dot Counting task along with the total correct and d-prime scores for the n-back task. Each task of EF includes aspects from multiple domains however for the purposes of illustrating the experimental procedures; the tasks are categorized accordingly below.

Fluency Tasks

Fluency refers to the ability to avoid response repetition while using strategies that maximize the formation of responses (Ruff, Allen et al. 1994).

Verbal Response Set

Verbal – Phonemic: Participants were instructed to say as many words as possible from a category in a given time (60 seconds). This category included 2 phonemic tasks, such as words that begin with letter L. Number of correct responses, repetitions, and rule violations were recorded.

Category – Animals: Participants were instructed to say as many words as possible from a category in a given time (60 seconds). This category included 2 semantic tasks, such as animals. Number of correct responses, repetitions, and rule violations were recorded.

Attention Tasks

Random Number: Participants were instructed to say numbers randomly from 0-9 until they reached 100 numbers. Total number of errors, strategies utilized, and prompts for randomness were recorded. This task was later dropped from the final battery for psychometric reasons.

Paper and Pencil Response Set

Design – Empty Dots: Participants were instructed to draw as many designs as possible with 4 straight lines, connecting empty dots only. They were told to complete this task as fast as they could, without touching any filled dots. The given time was 60 seconds. Rules for the second task were then modified such that participants were instructed to draw as many designs as possible alternating between connecting empty and filled dots in 60 seconds. They were also told to complete this task as fast as they could, without touching any filled or unfilled dots outside their designs. This task was dropped later from the final battery for psychometric reasons.

Planning, Problem Solving, and Abstract Reasoning Task

Planning is the ability to divide an objective into the individual components necessary to achieve the goal (Lezak, Howieson et al. 2004). Therefore effective planning requires a multitude of simultaneous processes, including sustained attention, abstract thinking, temporal sequencing, problem solving and reasoning (Norman and Shallice, 1986).

Unstructured Task: Participants were given 6 minutes to earn as many points as possible by accurately completing puzzles. Each puzzle was worth a specific amount of points, which was randomly associated with difficulty levels. They were given 3 booklets of puzzles and each booklet contained an equal sum of points. The number of high and low value puzzles that were accurately completed and the total number of points earned was recorded. This task was later dropped from the final battery for psychometric reasons.

Computer Response Set

Working Memory and Inhibition Tasks

Working memory is the ability to actively hold numerous pieces of information, for the execution of both verbal and nonverbal tasks, and subsequently retrieve them for further information-processing (Becker & Morris, 1999). On the other hand, inhibition is the ability to deny a response or suppress irrelevant or intrusive stimuli. Inhibition has been shown to have both "cognitive" and "behavioral" components (Chamberlain, Blackwell et al. 2005).

Flanker: Participants were shown a series of arrows, pointing to the left or right. They were instructed to press the left arrow key if the central arrow points to the left and

press the right arrow key if the central arrow pointed to the right. The target was flanked by arrows which correspond either to the same directional response as the target or to the opposite response. Response times for congruent and incongruent stimuli are recorded.

Dot Counting: Participants were shown a series of images containing blue circles, green circles, and blue squares. They were instructed to count the number of blue circles out loud on each screen, repeat the final number out loud, and remember that number. After a number of displays they were presented with screen containing question marks and expected to repeat the final numbers counted for each screen.

Set Shifting

Set shifting is the ability to concurrently or alternately engage in two different tasks (Crossley, Hiscock et al. 2004).

Set Shifting Task: Participants were shown a picture in the middle of the screen, and a word at the bottom of the screen. They were instructed to match the picture by shape or color with the objects in the corners of the screen. The word at the bottom of the screen told the subjects how to match the picture in the middle of the screen. When they matched the picture by color, they were instructed to push the left arrow key. When they matched the picture by shape they were instructed to push the right arrow key. Response times for accurate and inaccurate responses were recorded.

Attention

Attention is the ability to process selected information to the exclusion of others, in the service of achieving a goal. (Cohen, 2004)

Continuous Performance Task/Go-No-Go: Participants were presented with a number of shapes on the screen, one at a time. If a 5 pointed star appeared, they were

instructed to press the left arrow key. If any other shape appeared, they were instructed to not press any key. Response times for accurate and inaccurate responses were recorded.

1-Back: Participants were instructed to remember the location of the square that appeared on the screen, so that they could compare that location with the location of the next square that appeared on the screen. If the location was the same as the square before, they were told to push the left arrow key. If the location was different than the location of the square before, they were told to press the right arrow key. Response times for accurate and inaccurate responses were recorded.

Saccades (Go-No-Go): Participants were instructed to look at the center of the computer screen. They were instructed to move only their eyes in the direction that the dot moves on the screen or the opposite direction that the dot moves on the screen, depending on the trial. The total correct initial eye movements were recorded.

Final EXAMINER Scores

Using a combination of IRT methods and confirmatory factor analytic methods the final EXAMINER structure yielded both a one-factor and a three-factor organization that both showed strong support from factor analysis (Kramer et al., 2014). Eleven key variables from the battery (only 10 of these were used for children) were used to generate an overall Executive Function Composite score based on all of these variables and three specific factors: Fluency (based on total scores for Verbal Fluency and Category Fluency tests), Cognitive Control (based on four scores from the Flanker Test, Set Shifting test, anti-saccade test, and total dysexecutive errors), and Working Memory (based on four scores from N-back tests and the Dot Counting test in adults and three scores in children due to difficulties with 2-back version of the N-back test).

Given the use of IRT to measure factors with fewer items, test-retest reliability was chosen as the primary measure of reliability in the initial test development. The more traditional approach of internal consistency, assessed with Chronbach's alpha, is highly sensitive to the number of items in the scale. Therefore, test-retest reliability was assessed in 122 normal adult controls at an average interval of 25 days. The Executive Composite factor showed a test-retest reliability of r = .93. For the Fluency, Cognitive Control, and Working Memory factors the test-retest reliability was .88, .88, and .78, respectively.

The EXAMINER factor and composite scores represent the sums of weighted scores from the variables and are on an equal-interval scale. Each unit represents equal changes in ability across the scale. Scores are not age corrected.

Woodcock Johnson Tests of Cognitive Abilities 3rd edition

The WJ-III is designed to assess the CHC factors. It was normed on national sample of 4,783 children and 4,035 adults. The WJ-III was included in this investigation to assess verbal comprehension and knowledge, processing speed, and working memory as these measures have shown the most consistent and robust effects for SCD-related deficits. Picture Vocabulary, Synonyms, Antonyms, and Verbal Analogies are components within the Verbal Comprehension subtest. This subtest provides an assessment of comprehensive-knowledge. This subtest requires an individual to identify objects, provide knowledge of antonyms, and complete verbal analogies. Visual Matching is a subtest which provides a greater understanding of processing speed. Rapidly locating and circling identical numbers from a defined set of numbers, within a predetermined amount of time (3 min), is necessary for completion of this task. The

primary broad factor assessed by the Numbers Reversed subtest is Short Term Memory. The subject is instructed to hold a span of numbers in immediate awareness while verbally reversing the sequence. These subtests were given to create a greater understanding of the specific cognitive abilities for this population. The processing speed and short-term memory subtests are components of cognitive efficiency. Reliability coefficients indicate adequate to high reliability (.78 to .90) with children for all 3 subtests,

WJ-III tasks were scored according to the test manual. WJ-III raw scores and age-adjusted standard scores were mathematically transformed using the W scale, which is a data transformation used in clinical research and practice. The W scale score was created by converting raw scores such that they representing ability level with a center on a value of 500, which is set approximately at the average performance of children 10.0 years of age analysis (Woodcock et al., 2001). Each unit is meant to represent a similar change in ability level across the scale. The W metric from the WJ-III is most similar to the scaling of the EXAMINER factor scores and permitted the most direct comparison of both cognitive measures.

CHAPTER IV

STATISTICAL ANALYSIS PLAN

Hypothesis 1: The reliability (internal consistency) of the composite scores for EXAMINER will be equivalent to the WJ-III for use in pediatric SCD, including the reliability for youth tested across multiple sites, local demographic controls, and youth with SCD.

Internal consistency scores for the EXAMINER variables will be compared to the WJ-III measures by assessing the magnitude of difference between the two correlation values (Papoulis, 1990). WJ-III internal consistency scores will be for the 8-18 year olds in the normative sample who endorsed race as White or African American. Due to the construction of the final EXAMINER, the Executive Composite was used as the primary measure to examine the internal consistency of the measure across White and African-American participants. Due to differences in the age distribution between these two groups, twenty eight- or nine-year-old children were dropped from the White sample to create a better match of the two age distributions. The cases excluded were chosen at random using an on-line software program to select twenty cases from the total number of cases in this age range. In addition, there was 3.4% missing data for children included in the final analyses. I used mean substitution for missing data to allow for more cases to influence the reliability estimates. Imputation procedures are preferred in most cases for addressing missing data; however, in this context imputation would artificially increase

the internal consistency scores. All variables were centered with a mean of zero and standard deviation of one before computing internal consistency.

In the comparison of the internal consistency scores, the alpha value (which is essentially an intraclass correlation value among items) was compared using methods that compare two correlation values from separate samples. The two correlation coefficients are transformed with the Fisher r-to-z-transform and the probability for the magnitude of the difference between these two values can be computed (Lowry, 1998).

There are potential concerns about what is an adequate sample size for computing a stable internal consistency score. Many experts have recommended sample sizes of 200-300 as a minimal requirement (Peterson, 1994; Kline, 1986; Nunnally and Bernstein, 1994; Segall, 1994; Charter, 1999). Recent literature has indicated that the sample size needed for a stable measure of internal consistency is dependent on how unidimensional the items are in the scale (Yurdugul, 2008). This work provides estimates of sample sizes needed for stable measurement of alpha based on the magnitude of the first eigen value from principle components analysis of the scale items. Therefore, a principle component analysis was run with the healthy control sample to determine the first eigen value for the ten items used to compute the EXAMINER factor scores.

Hypothesis 2A: The strength of association between EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor and VIQ from the WJ-III will be similar for African American and White children.

Hypothesis 2B: The validity coefficient for the BRIEF Behavioral Regulation, Metacognition, and Global Executive Composite and the EXAMINER executive

composite, working memory factor, cognitive control factor, and fluency factor scores will be similar for African American and White children.

These hypotheses test the similarity in convergent validity across the two samples. The statistical tests will be calculated by comparing the magnitude of the two correlation values as described with *Hypothesis 1*, above.

Hypothesis 3A: Children with SCD at high risk for disease-related cognitive effects will show significantly poorer performance on the EXAMINER than demographically matched control children without SCD. The effect size of these differences will be compared to the magnitude of group differences on selected WJ-III.

Hypothesis 3B: The EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor will show at least as high of sensitivity and specificity in identifying high risk children with sickle-cell disease as the selected measures from the WJ-III relative to the local, demographically matched comparison group.

This aim will be evaluated by comparing the area under the curve using receiver operating characteristic (ROC) curves for each of the EXAMINER and WJ-III variables using the Statistical Package for Social Sciences software, version 22 (SPS-22). In previous work in SCD and other clinical conditions, sensitivity and specificity of measures has often been evaluated with cut-off scores that mimic the process often used in clinical practice, such as setting rules for the number of tests and the magnitude of difference (often 1.5 or 2.0 standard deviation units from age expectation) for indicating a positive test. However, I decided that this type of approach would be premature in working with the current version of the EXAMINER. The current EXAMINER version

does not contain age-adjusted scores and typically a large normative database would be used for establishing cut-off scores.

CHAPTER V

RESULTS

Hypothesis One: The reliability (internal consistency) of the EXAMINER will be equivalent to the WJ-III for use in pediatric SCD, including the reliability for youth tested across multiple sites, local demographic controls, and youth with SCD.

Descriptive Statistics

Prior to performing the statistical analyses, data were analyzed by SPSS-22 to detect the presence of outliers and the possible influence over variable relationships. Examination of histograms, independent samples t-tests, and chi-square analyses were performed to examine whether there were significant differences between the study groups. Normality was tested using the Shapiro – Wilk test. Overall, these procedures revealed no significant violations of normality, however a significant statistical age difference was found between the SCD group with White comparison in the final sample used, $t(362) = 4.94$, $p < .001$. The mean age of the SCD group ($M = 12.8$, $SD = 2.5$) was higher than White comparison group ($M = 11.2$, $SD = 2.9$), see Table 5.1. However, the overall age distribution (8-18 years), variance in age, and shapes of the distribution were highly similar across the two groups. Dropping additional cases to create a closer mean age was viewed as throwing away data and decreasing the reliability of the internal consistency estimates.

Internal Consistency

The reliability literature suggests that sample sizes should range from 200-500 to calculate a more precise estimate of coefficient alpha (Peterson, 1994; Kline, 1986; Nunnally and Bernstein, 1994; Segall, 1994; Charter, 1999). However, a recent investigation reveals that a sample size of 100 is adequate to produce a sufficiently unbiased estimator of coefficient alpha if the first eigenvalue is between 3.00 and 6.00 of the sample data set (Yurdugul, 2008). The principle components analysis of the 10 EXAMINER variables yielded an eigen value of 4.1 for the one-factor model, indicating the sample sizes in this study were likely to yield moderately to highly stable internal consistency estimates for this measure in the White and African-American samples. Using standard eigen value and scree plot rules, the principle components analysis also indicated a three-factor solution is appropriate, as found in the overall development of the EXAMINER battery, which used both exploratory and confirmatory factor analytic methods. The second and third factors had eigen values above 1 (1.3 and 1.2) and the scree plot showed a natural break in the curve with a relatively flat (horizontal) line after plotting the third factor.

Internal consistency of the EXAMINER and WJ-III were evaluated for Cronbach's alpha to assess the magnitude of difference between SCD and comparison groups. Internal consistency reliability values for 10 EXAMINER tasks were calculated to identify a consistent response pattern across the tasks, see Table 5.2. Although the EXAMINER contains 11 cognitive tests, the 2-back task was not relevant for this analysis. This sub-test of the N-back was only administered to the adult norming group. Results revealed the following Cronbach’s alpha scores for the EXAMINER: White

control group, α = .84 (*n* = 279), African American comparison group, *α* = .81 (*n* = 85), SCD group, *α* = .84 (*n* = 32). Reliability coefficient data for the WJ-III was obtained from Woodcock et al., (2001) using the mean internal consistency score: Verbal Comprehension, *α* = 0.91 (*n* = 3549, *SEM* = 4.40), Visual Matching, *α* = 0.88 (*n* = 3719, *SEM* = 5.15), and Numbers Reversed, *α* = 0.86 (*n* = 3337, *SEM* = 5.53).

Significance of Difference between Cronbach's Alpha Coefficients

Analyses were conducted to assess the significance of the difference between internal consistency coefficients. Chronbach's alpha scores for the SCD and comparison groups were then transformed into Fisher Z scores and the probability for the magnitude of the difference between the values was computed, with a two-sided test. These findings suggest statistically significant differences between reliability coefficients for the EXAMINER Composite values for White and Black youth on the WJ-III verbal comprehension and visual matching measures, see Table 5.3. Comparing reliability values from the EXAMINER to the WJ-III, for the White sample, reveal the following findings: verbal comprehension, (*z* = -4.9, *p* = .00), visual matching, (*z* = -2.48, *p* = .01), numbers reversed, (*z* = -1.15, *p* = -.25). Results for the WJ-III reliability scores compared to the EXAMINER composite, for the African-American sample, were as follows: verbal comprehension, (*z* = -3.59, *p* = .00), visual matching, (*z* = -2.23, *p* = .03), numbers reversed, (*z* = -1.49, *p* =.14). In addition, analyses of the alpha reliability data between EXAMINER composite and WJ-III reliability coefficients for the SCD group reveal no significant differences between values: verbal comprehension, (*z* = -1.64, *p* = .10), visual matching, (*z* = -0.83, *p* =.41), numbers reversed, (*z* = -0.39, *p* =.70).

For descriptive purposes, internal consistency scores were also computed for the specific EXAMINER factors, which indicated a generally similar pattern. There were not significant differences evident across race/ethnicity groups, as a test of potential cultural differences in the validity of the measures. However, the internal consistency values were lower than found for the WJ-III. The EXAMINER Fluency factor is based on the two fluency scores found in the EXAMINER composite variable. The alpha coefficient was .61 for the White sample, .65 for the African-American sample, and .68 for the SCD sample. The Working Memory factor was based on the variables from the Dot Counting and N-back tests (n = 3 variables) and showed an alpha of .76 for the White -sample, .72 for the African-American sample, and .63 for the SCD sample. Finally, the remaining variables are part of the EXAMINER Cognitive Control factor (n = 5 variables) and showed an alpha of .80 for the White sample, .78 for the African-American sample, and .71 for the SCD sample.

Hypothesis 2A: The strength of association between EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor and VIQ from the WJ-III will be similar for African American and White children.

Correlation Analysis

Partial correlations, controlling for age, were conducted to examine construct validity for EXAMINER in relation to verbal ability. Data were computed with pairwise exclusion for missing values to maximize the sample size for each validity coefficient

The verbal ability construct is represented by both the WRAT-IV reading T score and WJ-III verbal comprehension standard score. As mentioned above, the WJ-III was not administered to the White comparison group. For that reason, reading performance on

the WRAT-IV was included in these analyses. Inclusion of this sub-test is psychometrically suitable because the WRAT-IV reading score and WJ-III verbal comprehension have been shown to demonstrate a moderate-to-large (r = .48) Pearson correlation, at the $p > .05$ significance level (Woodcock et al., 2001).

Partial correlations were computed among the four scores of the EXAMINER and verbal ability, holding age constant. The results of these correlational analyses indicated that the EXAMINER executive composite, fluency factor, cognitive control factor, and working memory scores were positively correlated with verbal ability standard scores, see Table 5.4. These partial correlation values also revealed generally similar associations across the two groups, with a trend toward higher convergent validity for the African-American control sample as compared to the White control sample. The relationship between the EXAMINER composite scores and verbal ability for the White comparison group were small-to-medium in size (r = .29 to .42, $p < .05$). Similarly, the strength of the association between the EXAMINER and verbal ability for the African American comparison group were medium-to-large associations (r = .28 to .50, $p < .05$). Descriptive terms are based on those used by Cohen (1988).

Significance of Difference between Correlation Coefficients

Fisher r-to-z transformations were conducted, across the two samples, to assess the significance of the difference between partial correlation coefficients which evaluated construct validity for the EXAMINER. Partial correlation scores for the cognitive performance, controlling for age, of the African American control and White comparison groups were transformed into Fisher Z scores and the probability for the magnitude of the difference between the values was computed, with a two-sided test. The strength of

association between the EXAMINER executive composite (z = -.72, p < .05), fluency factor (z = -.44, p < .05), cognitive control (z = -.38, p < .05), and working memory (z = -.01, p < .05) and verbal ability was not statistically significant for African American and White samples.

Hypothesis 2B: The validity coefficient for the BRIEF Behavioral Regulation, Metacognition, and Global Executive Composite and the EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor scores will be similar for African American and White children.

Correlation Analysis

Partial correlations were computed to examine construct validity for EXAMINER compared to the BRIEF, controlling for age. To maximize sample size for each reliability coefficient, data were computed with pairwise exclusion for missing values.

The degree of association between EXAMINER executive composite, fluency factor, cognitive control, and working memory scores and the BRIEF Metacognition, Behavioral Regulation, and Executive Composite scales was examined. Results of these correlational analyses indicated that all EXAMINER variables were correlated in the expected direction with each of the BRIEF indices, see Table 5.4. For the EXAMINER composite, fluency, and working memory factors higher scores mean better performance. For the EXAMINER cognitive control factor higher scores mean more difficulties with cognitive control tasks. These partial correlation values also suggested a trend of higher validity coefficients for the African-American control sample as compared to the White control sample for the EXAMINER. The relationship between the EXAMINER scores and BRIEF Metacognitive Index, controlling for age, were small for the White

comparison group (r = -.14 to -.21, p < .05) and small-to-medium in size for the African American control group (r = -.18 to -.30, p < .05). The relationship between the EXAMINER scores and BRIEF Behavioral Regulation Index, holding age constant, also ranged from very small-to-medium range for the White comparison group (r = -.01 to -.21, p < .06) yet the range of scores for the African American control group were small-to-medium in size (r = -.18 to -.30, p < .05). Finally, the relationship between the EXAMINER scores and BRIEF Executive Composite Index, holding age constant, were small for the White comparison group (r = -.12 to -.17, p < .05) but medium-to-large in the African American control group (r = -.25 to -.32, p < .05).

Significance of Difference between Correlation Coefficients

Fisher r-to-z transformations were conducted to assess the significance of the difference between partial correlation coefficients utilized in EXAMINER construct validity evaluation. Partial correlation scores assessing the cognitive performance of African American control and White comparison groups between BRIEF and EXAMINER factors were transformed into Fisher Z scores.

The probability for the magnitude of the difference between these values was computed across the two samples, with a two-sided test. Ten calculations for the test of the difference between EXAMINER and BRIEF scores were found to be statistically non-significant, see Table 5.4. The strength of association between EXAMINER executive composite (z = -.62, p < .05), fluency factor (z = .83, p < .05), cognitive control (z = -.02, p < .05), and working memory factor (z = .86, p < .05) and the BRIEF Metacognition Index was not statistically significant between the African American and White samples. The strength of association between EXAMINER fluency factor (z =

1.57, $p < .05$) and working memory factor ($z = 1.27$, $p < .05$) and the BRIEF Behavioral Regulation Index was not statistically significant between the African American and White samples. Conversely, the strength of association between EXAMINER executive composite ($z = 2.36$, $p < .05$) and cognitive control factor ($z = 2.61$, $p < .01$) in relation to the BRIEF Metacognition were significantly larger for the African American than the White sample. The strength of association between EXAMINER executive composite ($z = 1.16$, $p < .05$), fluency factor ($z = 1.11$, $p < .05$), cognitive control ($z = -.89$, $p < .05$), and working memory factor ($z = .86$, $p < .05$) and the BRIEF Executive Composite Index was not statistically significant between the African American and White samples.

Hypothesis 3A: Children with SCD at high risk for disease-related cognitive effects will show significantly poorer performance on the EXAMINER than demographically matched control children without SCD. The effect size of these differences will be compared to the magnitude of group differences on selected WJ-III.

Descriptive Statistics

Prior to performing the statistical analyses, data were analyzed by SPSS-22 for the presence of outliers and the possible influence over variable relationships. Examination of histograms and a series of independent sample t-tests were performed to examine whether there were significant differences between the age ranges of SCD and comparison study group. Normality was tested using the Shapiro – Wilk test. Overall, these procedures revealed no significant violations of normality. These procedures revealed no significant statistical difference in age between the SCD group ($M = 12.72$, $SD = 3.16$) and the African American comparison group ($M = 12.87$, $SD = 2.53$), see Table 5.1.

Analysis of Variance Analyses

An ANOVA analysis revealed a statistically significant difference between the parental income levels of the SCD group and the African American comparison group, at the p<.05 level [F (21.110) = 20.33, p = .00, η^2 = .01]. The mean level of parental income for the African American sample was (M = 2.92, SD= 1.22) was higher than the mean level of parental income for the SCD group (M = 1.74, SD= 1.40). Therefore income was added to the statistical models to control for this variance.

A series of two-way analysis of variance analyses were conducted to examine whether there were statistically significant differences among children with SCD and the non-neurologic control group, in relation to performance on the EXAMINER.

ANOVA results revealed a main effect between the SCD group and African American comparison group on the executive composite score, at the p<.05 level [F (1, 111) = 12.86, p = .00, η^2 = .13]. The main effect of income was non-significant at the p<.05 level [F (4, 111) = 1.37, p = .25, η^2 = .05]. The average executive composite score was significantly higher for African American control children (M = -.18, SE = .08) than the SCD group (M = -.72, SE = .10) composite scores, see Figure 5.1. This evidence suggests that children diagnosed with SCD have poorer performance on the EXAMINER when compared to demographically matched peers.

ANOVA results also reveal that the main effect of the SCD group and the African American comparison group, in relation to the cognitive control factor, was statistically significant, controlling for family income levels, at the p < .05 level [F (1, 111) = 12.66, p = .00, η^2 = .11]. The main effect of income was non-significant at the p<.05 level [F (4, 107) = 1.26, p = .29, η^2 = .04]. The mean composite cognitive control scores for the SCD

group (M = -.60, SE = .14) were significantly lower than the Black comparison group (M = .06, SE = .11) cognitive control scores, see Figure 5.2. Results suggest that children with SCD exhibit poorer cognitive control than Black peers.

ANOVA analysis also examined the presence of statistically significant differences among children, controlling for family income levels, diagnosed with SCD and a demographically matched comparison group related to their performance on the EXAMINER fluency factor score. Results revealed statistically significant differences among the groups, at the $p < .05$ level [F (2, 393) =8.08, p = .00, η^2 = .04]. The main effect of income was non-significant at the $p<.05$ level [F (4, 105) = .47, p = .76, η^2 = .02]. The mean fluency scores of the SCD group (M = -.68, SE = .10) were statistically lower than those of the demographically matched control group (M = -.20, SE = .08), see Figure 5.3. Findings indicate that African American children with no neurologic condition performed significantly higher compared to children diagnosed with SCD on the mean fluency composite score.

ANOVA findings also demonstrate a significant main effect between SCD group status and the African American comparison group on the EXAMINER working memory composite score, at the $p<.05$ level [F (1, 92) = 7.18, p = .01, η^2 = .17]. The main effect of income was non-significant at the $p < .05$ level [F (4, 92) = 1.28, p = .28, η^2 = .05].The mean score for the African American comparison group (M = -.21, SE = .09) was significantly different than the SCD group (M = -.64, SD = .13), see Figure 5.4. Results suggest that African American children with no neurologic condition performed better than children diagnosed with SCD on the working memory factor score, η^2 = .17].

Similarly, ANOVA results revealed statistically significant differences between the SCD group and the demographically matched control group, across income levels, on the W scores calculated from the WJ-III. Results revealed that the verbal comprehension scores for the SCD group and the African American comparison group were statistically different, at the p<.05 level [F (1, 111) = 7.12, p = .01, η^2 = .07]. The main effect of income was non-significant at the p<.05 level [F (4, 111) = .84, p = .50, η^2 = .03].Verbal comprehension scores for the SCD group (M = 495.60, SE = 2.59) was significantly lower than the Black comparison group (M = 504.94, SE = 2.01) verbal comprehension scores, see Figure 5.5. This evidence suggests that children diagnosed with SCD have poorer verbal comprehension on the WJ-III when compared to demographically matched peers.

In addition, ANOVA analysis examined the presence of statistically significant differences among children with SCD and a demographically matched comparison group in relation to their performance on WJ-III visual matching score, across all income levels. Results revealed a statistically significant main effect among the groups, at the p<.05 level [F (1, 111) = 14.35, p = .00, η^2 = .12]. The main effect of income was non-significant at the p<.05 level [F (4, 111) = .89, p = .47, η^2 = .03]. There were statistically significant differences between the SCD group (M = 495.88, SE = 2.75), and those of the demographically matched control group (M = 509.38, SE = 2.13), see Figure 5.6. Findings indicate that African American children with no neurologic condition performed significantly higher compared to children diagnosed with SCD for visual matching on the WJ-III.

ANOVA results also revealed a significant main effect of SCD group status on WJ-III numbers reversed W score across income levels, at the p<.05 level [F (1, 111) = 7.21, p = .00, η^2 = .06]. The main effect of income was non-significant at the p<.05 level [F (4, 111) = 1.09, p = .37, η^2 = .04]. The mean score for the African American comparison group (M = 504.49, SE = 3.67) was significantly different than the SCD group (M = 487.99, SD = 4.73), see Figure 5.7. Results suggest that African American children with no neurologic condition performed better than children diagnosed with SCD, on the WJ-III numbers reversed task.

The eta-squared values for the EXAMINER scores ranged from small-to-large, .04 to .17 with a median of .12, whereas for the WJ-III the effect sizes were medium, .06 to .12 with a median of .07. Comparison of the magnitude of the observed effects across the EXAMINER and WJ-III W scale measures might suggest a slight trend toward larger effects for the EXAMINER, though overall the values were comparable (Cohen, 1988).

Overall, children with SCD showed significantly poorer performance on the EXAMINER than demographically matched control children without SCD, Table 5.5. Therefore, hypothesis 3A was supported. These results further support the conclusion that children with SCD evidence disease-related cognitive effects.

Hypothesis 3B: The EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor will show at least as high of sensitivity and specificity in identifying high risk children with sickle-cell disease as the selected measures from the WJ-III relative to the local, demographically matched comparison group.

Descriptive Statistics

Prior to performing the statistical analyses, SPSS-22 was utilized to conduct a subgroup analysis of African American comparison group. This procedure was performed to provide balance in parental income between groups. A targeted selection of cases was performed from the African American comparison group, with the intent of exclude cases more than 1 *SD* away from mean income level of the SCD group (*M* = 1.73, *SD* = 1.74). A random sub-group (*n* = 51) was then generated with a lower mean parental income level (*M* = 1.39, *SD* = .73). Normality was tested using the Shapiro–Wilk test. Analyses revealed no significant violations of normality or significant statistical difference between the SCD group (*M* = 12.72, *SD* = 3.16) with skewness of 0.8 (SE = 0.41) and kurtosis of -1.2 (SE = 0.8) and the African American comparison group (*M* = 12.64, *SD*= 2.77) with skewness of 0.01 (SE = .44) and kurtosis of -0.88 (SE = 0.86). An independent samples t-test revealed no statistical significance between the income range of the SCD and African American comparison groups (*t* = 9.99, *df* = 27, p < .73). ROC Curves were constructed based upon the sensitivity and specificity of the EXAMINER and WJ-III scores.

Receiver Operating Characteristic Curve Analysis

ROC Curve analyses were conducted to evaluate the performance of multiple curves in order to compare the EXAMINER and WJ-II neuropsychological measures. Results indicated that the EXAMINER scores were statistically significance (*p* = .05) in differentiating between the SCD and demographically matched comparison groups. The balance between sensitivity and false positive rate is shown in Figure 5.8. The WJ-III

scores were also statistically significance (p = .05) for differentiating between study groups. The balance between sensitivity and false positive rate is shown in Figure 5.9.

ROC Curves in Figure 5.8 and Figure 5.9 can be expressed numerically as area under the curve values, which indicate the likelihood that a randomly selected case will perform more poorly than a randomly selected control. The area under the curves for the EXAMINER scores were: executive composite (.79), fluency factor (.74), cognitive control factor (.77), working memory factor (.74) and these values appear to be similar to the WJ-III: verbal comprehension, (.72), visual matching (.76), numbers reversed (.70). More specifically, for discriminating no SCD status from SCD diagnosis among all subjects, the areas under the ROC curves for the EXAMINER and WJ-III evidence moderate accuracy by Swets's (1988) criteria, see Table 5.6.

Table 5.1.

Descriptive Information for Study Groups

Variable	High Risk SCD (n-32)	AA Comparison Group Without SCD (n-85)	White Comparison Group Without SCD (n-279)	Statistic
Age (*M* ± *SD*; range)	12.7 ± 3.2	12.8 ± 2.5	11.2 ± 2.9	*t*(362) = 4.94, *p* < .001
Gender (*n*)				
Males	14 (43.8%)	39 (46.3%)	133 (47.0%)	
Females	18 (56.3%)	46 (59.7%)	146 (53.0%)	
Race/Ethnicity				
African-American	32	80 (94.5)	n/a	
Hispanic and African American	0	5 (0.6%)	n/a	
Parent Education				
9-11 years	3 (3.2%)	4 (3.0%)	n/a	
High School	4 (13.3%)	19 (16.2%)	n/a	
Some College	9 (28.8%)	20 (17.0%)	n/a	
College Degree	16 (51.2%)	42 (35.7%)	n/a	
Family Income (*n*)				*F(21.110) = 20.33, p = .001*
< $10,000/year	9 (28.8%)	3 (3.5%)	n/a	
$10-20,000/year	5 (15.6%)	9 (10.6%	n/a	
$20-30,000/year	8 (25.0%)	15 (17.6%)	n/a	
$30-40,000/year	5 (15.6%)	18 (21.2%)	n/a	
> $40.000/year	5 (15.6%)	40 (47.1%)	n/a	
Adults in the home (*M* ± *SD*)	1.6 ± 0.5	1.8 ± 0.6	n/a	
Age of Primary Caregiver (*M* ± *SD*)	37.4 ± 9.0	37.0 ± 9.0	n/a	
Current Therapy (*n*)				
Hydroxyurea	2	n/a	n/a	
Chronic therapy	1	n/a	n/a	
Bone Marrow therapy	0	n/a	n/a	
Routine CBC Values				
White cells (K/ul)	13.1 ± 3.1	n/a	n/a	
Hemoglobin (g/dL)	8.6 ± 1.3	n/a	n/a	
Platelets (K/uL)	466 ± 129	n/a	n/a	

SCD=sickle cell disease
*p<0.05; **p<0.01

Table 5.2.

Overview of Internal Consistency for EXAMINER and WJ-III

Test	Study Groups			WJ-III Normative Sample
	African American Comparison Group	White Comparison Group	SCD Group	
Examiner Battery				
Composite Score	.81 (*n* = 85)	.84 (*n* = 279)	.84 (*n* = 32)	
Woodcock-Johnson III				
Verbal Comprehension	-	-	-	.91 (*n* = 3549)
Visual Matching	-	-	-	.88 (*n* = 3719)
Numbers Reversed	-	-	-	.86 (*n* = 3337)

Note: Internal consistency reliability values, Chronbach's alpha, for 10 EXAMINER tasks and Woodcock-Johnson Tests of Cognitive Abilities, 3rd edition mean internal consistency scores for children 8-18 of age; *p<.05; **p<.01,2-tailed test.

Table 5.3.

Fisher r-to-z transformation scores used to assess the significance of the difference between EXAMINER composite scores and WJ-III tests.

WJ-III Subscales	African American	Sickle Cell Disease	White
Verbal Comprehension	-3.59**	-1.64	-4.9**
Visual Matching	-2.23*	-0.83	-2.48**
Numbers Reversed	-1.49	-0.39	-1.15

Notes: Z-scores test the difference in the magnitude of the correlation between the Woodcock-Johnson Tests of Cognitive Abilites, 3rd edition Verbal Comprehension, Visual Matching, and Numbers Reversed scales and the EXAMINER Composite scores for the African-American, SCD, and White samples; *p<.05; p<.01, 2-tailed test.

Table 5.4.

Partial correlations (controlling for age) assessing construct validity for EXAMINER Battery factor scores. Twelve of sixteen coefficients were in the direction of higher convergent validity for the African-American sample as compared to the White sample.

Measure	Group	Executive Composite	Fluency	Cognitive Control	Working Memory
Verbal Ability – Age Adjusted Score	W	.421 (n = 231)	.293 (n = 232)	-.380 (n = 231)	.279 (n = 218)
	A-A	.495 (n = 82)	.345 (n = 82)	-.339 (n = 79)	.280 (n = 73)
		z = -0.72	z = -0.44	z = 0.36	z = -0.01
BRIEF - Meta-cognitive Index	W	-.212 (n = 246)	-.142 (n = 246)	.187 (n = 246)	-.180 (n = 231)
	A-A	-.291 (n = 72)	-.250 (n = 73)	-.184 (n = 70)	-.297 (n = 64)
		z = 0.62	z = 0.83	z = -0.02	z = 0.86
BRIEF - Behavioral Regulation Index	W	-.054 (n = 246)	-.059 (n = 246)	.006 (n = 246)	-0.153 (n = 231)
	A-A	-.352 (n = 77)	-.261 (n = 77)	.344 (n = 74)	-.228 (n = 68)
		z = 2.36*	z = 1.57	z = 2.61**	z = 1.27
BRIEF – Executive Composite Index	W	-.170 (n = 244)	-.123 (n = 244)	.130 (n = 244)	-.145 (n = 229)
	A-A	-.318 (n = 72)	-.269 (n = 72)	.249 (n = 69)	-.273 (n = 63)
		z = 1.16	z = 1.11	z = 0.89	z = 0.92

Notes: Verbal Ability was estimated from either the Woodcock-Johnson Tests of Cognitive Abilities, 3rd edition Verbal Comprehension scale or the Wide Range Achievement Test, 4th edition Reading Score; BRIEF=Behavioral Rating Inventory of Executive Function; Z-scores test the difference in magnitude of the correlation between the White and African-American samples; *p<.05; **p<.01; data were computed with pairwise exclusion for missing values to maximize the sample size for each validity coefficient.

Table 5.5.

Mean Score for African American and SCD study groups on EXAMINER Z scores and WJ-III W Scores

	Study Groups	
	African American Comparison Group	SCD Group
Test		
Examiner Battery		
Composite Factor Score	-0.18 (*SE* = .08)	-0.72 (*SE* = .10)
Cognitive Control Factor Score	-0.60 (*SE* = .14)	0.06 (*SE* = .11)
Fluency Factor Score	-0.20 (*SE* = .08)	-0.68 (*SE* = .10)
Working Memory Factor Score	-0.21 (*SE* = .09)	-0.64 (*SE* = .13)
Woodcock-Johnson III		
Verbal Comprehension	504.94 (*SE* = 2.01)	495.60 (*SE* = 2.59)
Visual Matching	509.38 (*SE* = 2.13)	495.88 (*SE* = 2.75)
Numbers Reversed	504.49 (*SE* = 3.67)	487.99 (*SE* = 4.73)

Note: Internal consistency reliability values, Chronbach's alpha, for 10 EXAMINER tasks and Wookcock-Johnson Tests of Cognitive Abilities, 3rd edition mean internal consistency scores and standard error scores for children 8-18, *p<.05; **p<.01, 2-tailed test.

Table 5.6

Area under ROC curve for EXAMINER and WJ-III

	Area Under the Curve (95% CI)	Standard Error
EXAMINER		
Composite Factor Score	.79* (.67-.90)	.06
Cognitive Control Factor Score	.77* (.66-.88)	.06
Fluency Factor Score	.74* (.62-.86)	.06
Working Memory Factor Score	.74* (.63-.86)	.06
Woodcock-Johnson III		
Verbal Comprehension W Score	.72* (.61-.82)	.06
Visual Matching W Score	.76* (.65-.87)	.06
Numbers Reversed W Score	.70* (.59-.80)-	.05

Note: Null hypothesis: true area = 0.5
*. Significant at the 0.05 level (2-tailed).

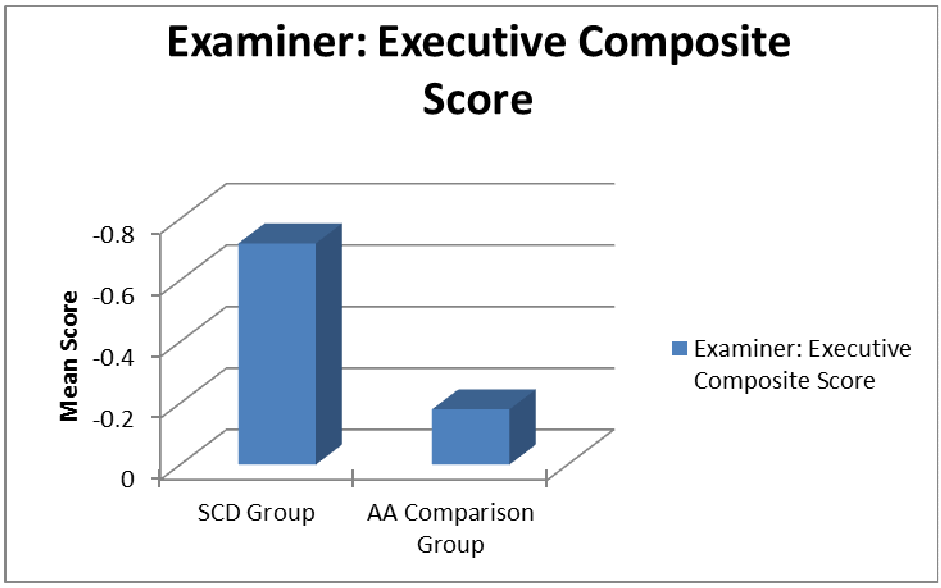

Figure 5.1. Mean EXAMINER Composite Scores for SCD and AA Comparison Group

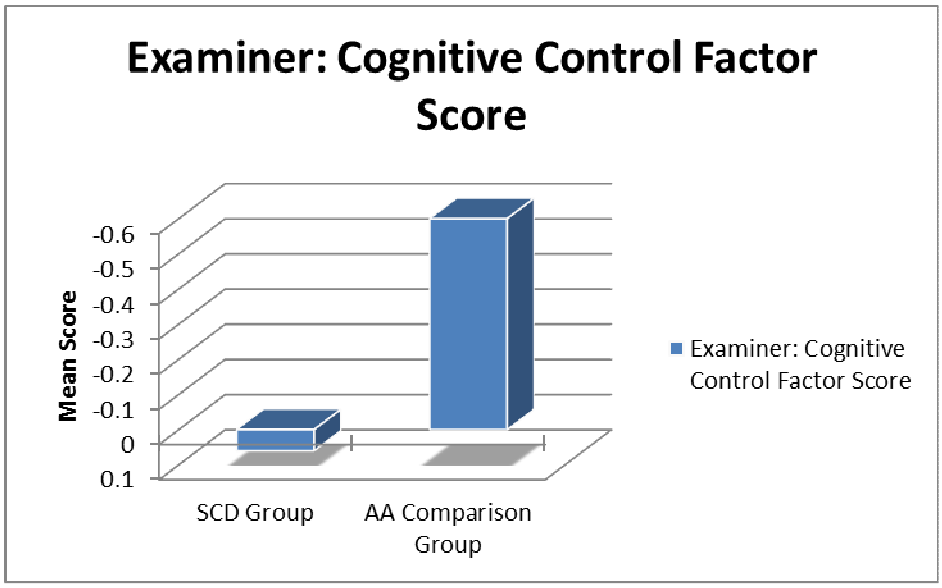

Figure 5.2. Mean EXAMINER Cognitive Control Scores for SCD and AA Comparison Group

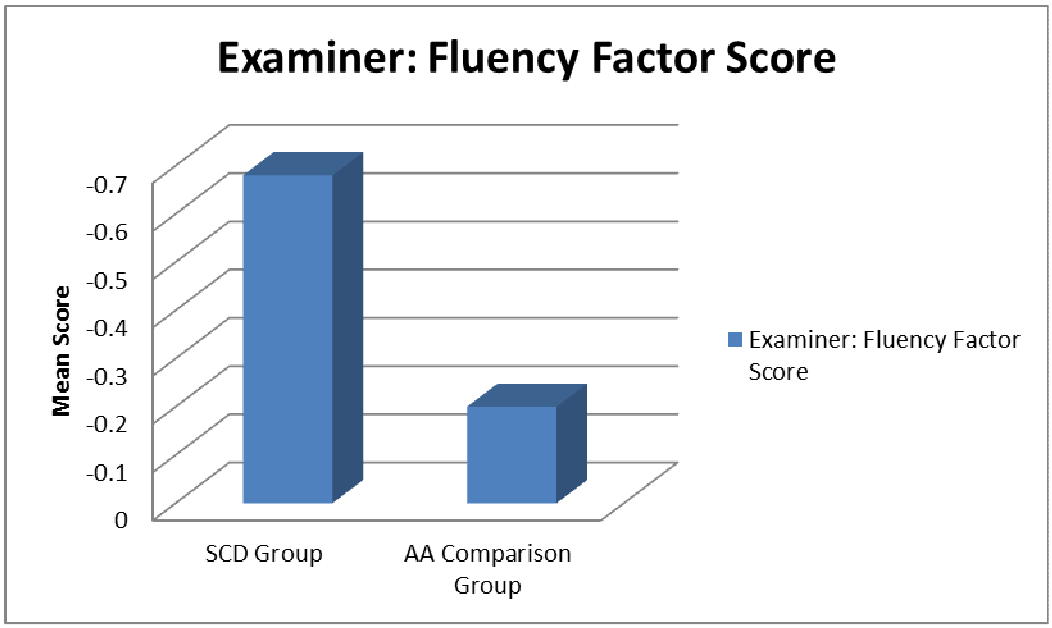

Table 5.3. Mean EXAMINER Fluency Scores for SCD and AA Comparison Group

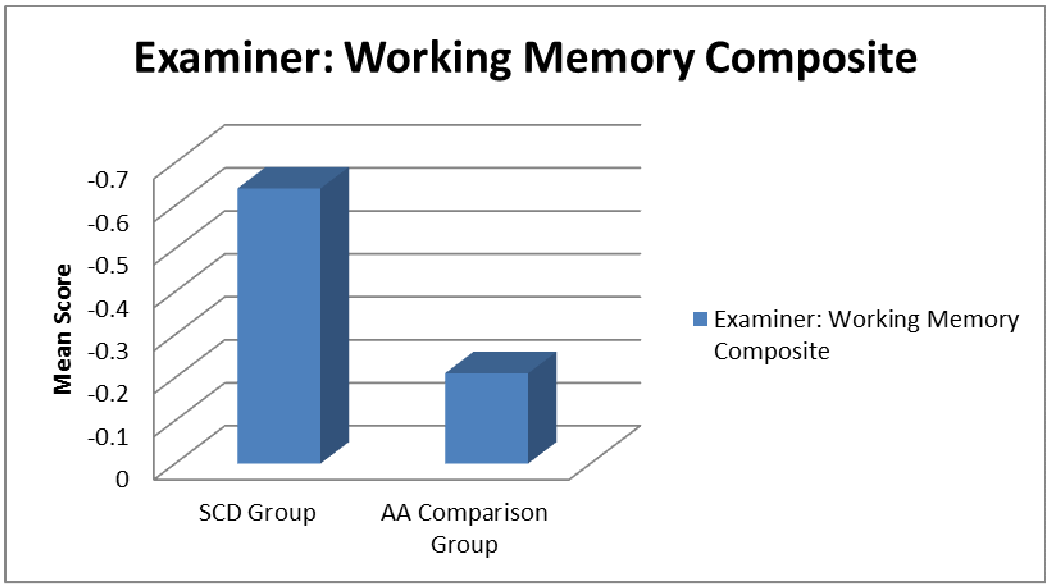

Figure 5.4. Mean EXAMINER Working Memory Scores for SCD and AA Comparison Group

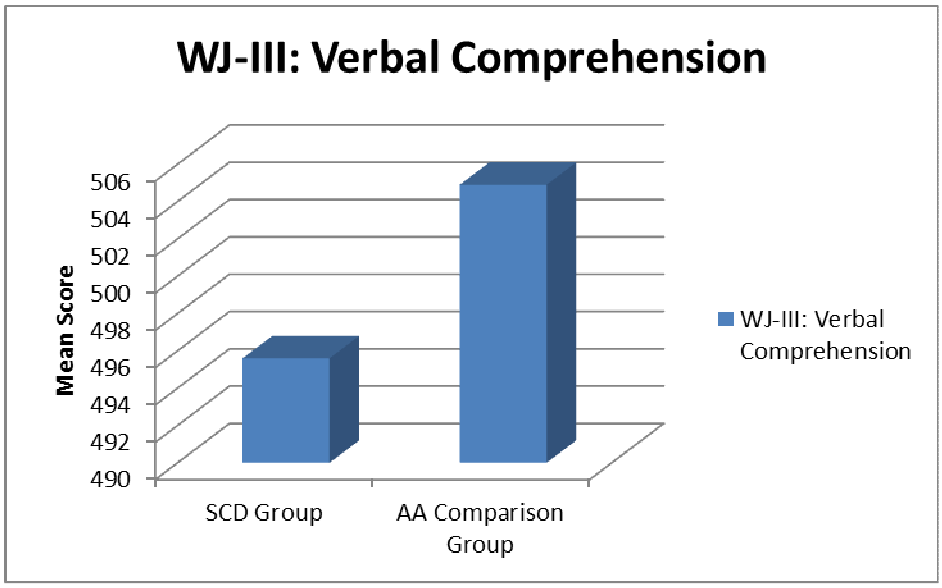

Figure 5.5. Mean WJ-III Verbal Comprehension Scores for SCD and AA Comparison Group

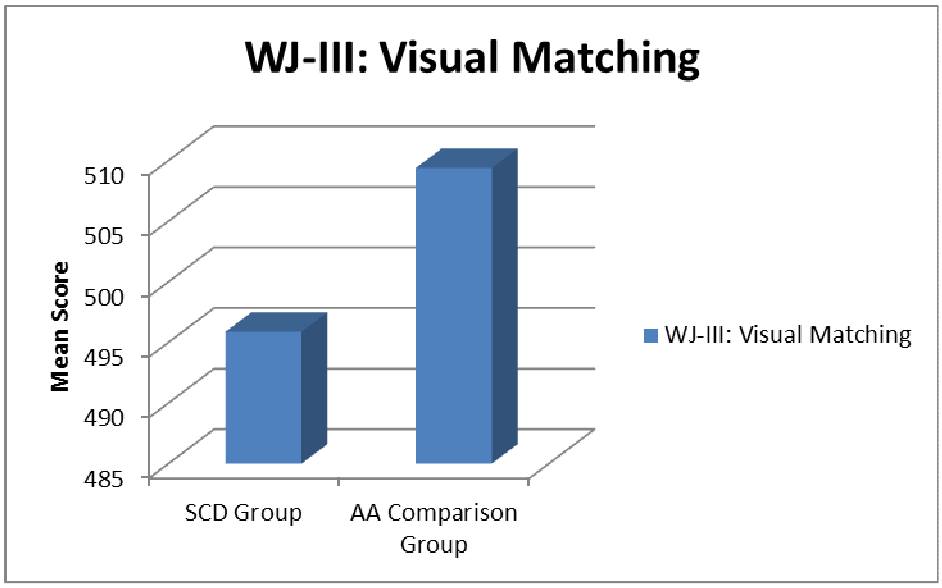

Figure 6.6. Mean WJ-III Visual Matching Scores for SCD and AA Comparison Group

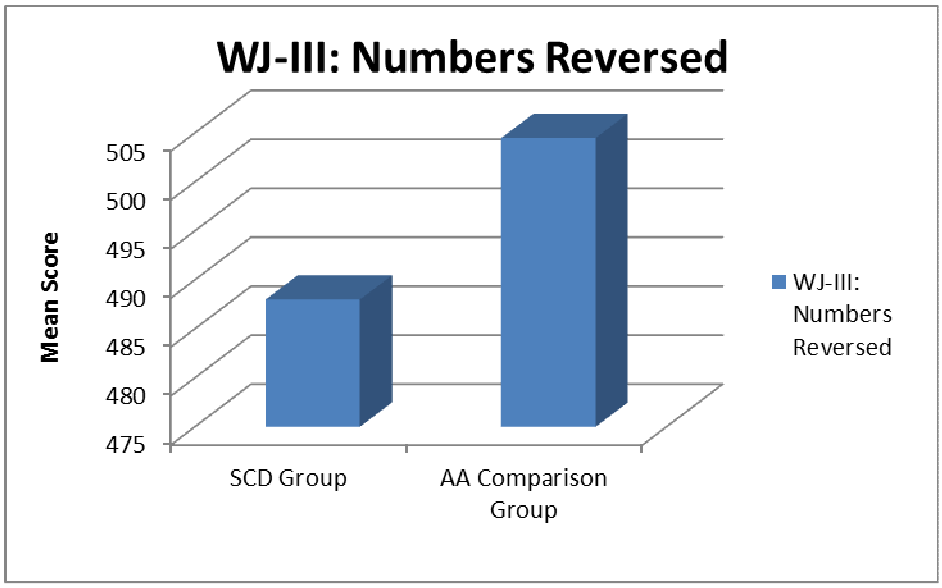

Figure 5.7. Mean WJ-III Numbers Reversed Scores for SCD and AA Comparison Group

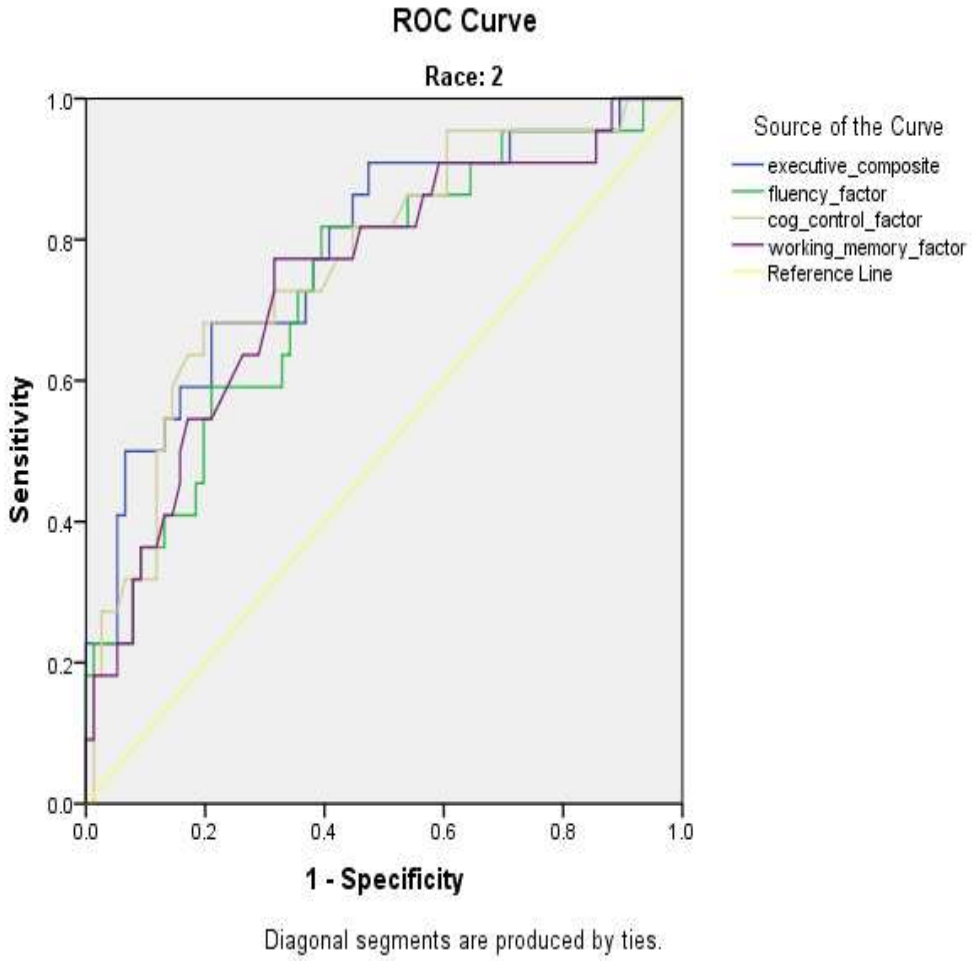

Figure 5.8. Hypothesis 3B - Receiver-Operating Characteristic Curves for EXAMINER Factor and Composite Scores

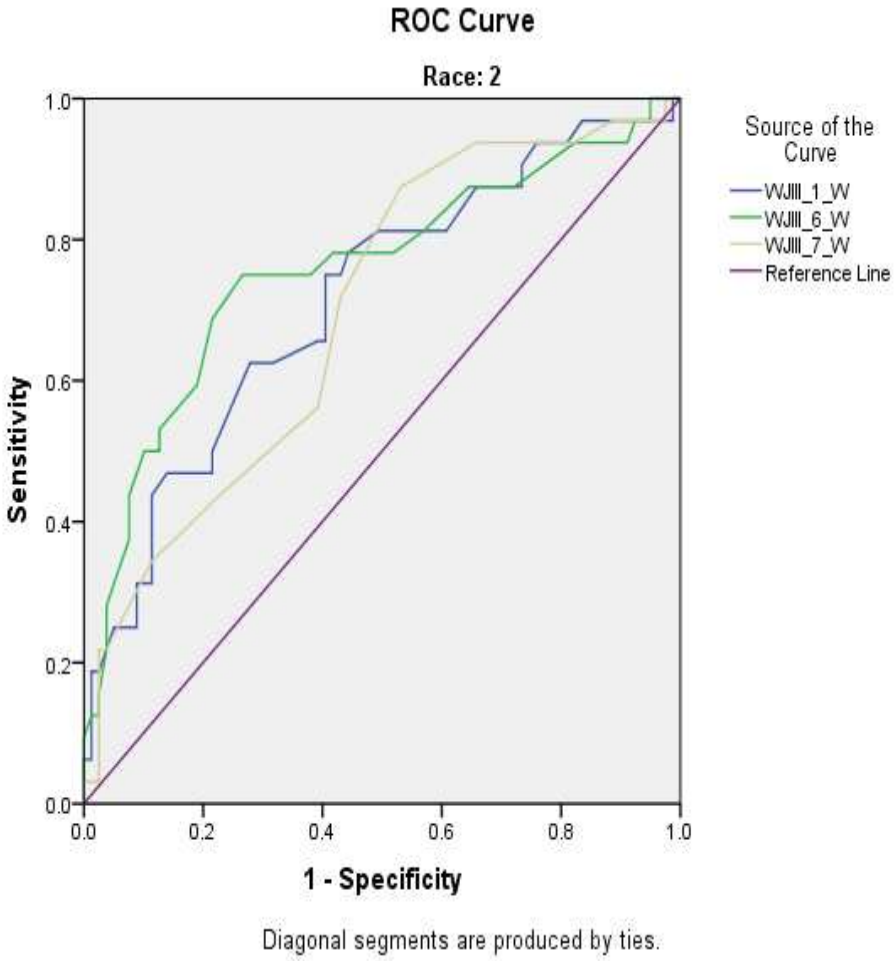

Figure 5.9. Hypothesis 3B - Receiver-Operating Characteristic Curves for WJ-III Verbal Comprehension, Visual Matching, and Numbers Reversed tests

CHAPTER VI

DISCUSSION

This study aimed to investigate the psychometric properties of a new executive function measure, the EXAMINER, to determine its utility with the pediatric SCD population. The psychometric properties of the EXAMINER were evaluated, relative to the WJ-III. The WJ-III was chosen as a gold standard because this literature review revealed that it has strong psychometric properties and cultural validity with African-American children. Thus, the WJ-III appears to be the most appropriate measure for use with the pediatric SCD population, in the US. In this investigation, the cultural validity of the EXAMINER was investigated for African-Americans with SCD, by comparing the internal consistency and convergent validity for African-American youth relative to White youth. In addition, criterion validity was assessed for the EXAMINER to SCD-related cognitive deficits. Overall, this study revealed three novel findings. First, the EXAMINER demonstrated adequate reliability (internal consistency) for the Executive Composite score. This assessment of internal consistency was based on traditional standards. Reliability values were found to be similar across White and African-American children. However, reliability was found to be lower for the EXAMINER, compared to the normative sample of children in the WJ-III manual. Second, the EXAMINER displayed acceptable cultural validity for African-Americans with SCD. Convergent validity was evaluated by the degree to which the EXAMINER and well established measures were related across Black the White samples. Finally, the

children with and without neurologic morbidity and performed at least as well as an established measure, the WJ-III.

Hypothesis 1: The reliability (internal consistency) of the EXAMINER will be equivalent to the WJ-III for use in pediatric SCD, including the reliability for youth tested across multiple sites, local demographic controls, and youth with SCD.

Results suggested that hypothesis one was partially supported by the findings in this investigation. Findings revealed that the internal consistency of the EXAMINER composite score appeared to function adequately across SCD and comparison pediatric groups. Similar to the WJ-III, the EXAMINER composite score evidenced good internal consistency for each study group when employing commonly accepted descriptive labels (Cohen, 1988). More specifically, the internal consistency of the EXAMINER for each of the study groups reached the generally accepted 0.8 Chronbach's alpha threshold for making interpretations and clinical predictions. There were also no significant differences for internal reliability found between the pediatric SCD, African American, and White comparison groups on the EXAMINER, relative to the WJ-III. Likewise, there were no differences found within the pattern of responses between the EXAMINER and WJ-III for children diagnosed with pediatric SCD and youth in the comparison groups for the WJ-III working memory task. Divergent to Hypothesis 1, the magnitude of difference between the internal consistency of the EXAMINER composite score and WJ-III verbal comprehension and visual matching tests were significant for White and Black youth, indicating higher reliability for these WJ-III measures than for the EXAMINER.

Contrary to what we originally expected, the EXAMINER showed lower reliability than found in the WJ-III manual for the measure's normative sample. It is

possible that comparison between a larger and more diverse sample, matched to the U.S. census, improved the alpha value by increasing the range of scores with greater variability. However, the EXAMINER appeared to perform nearly as well as the WJ-III with one-third the items. This suggests the EXAMINER items were well chosen. The WJ-III measures typically have 30 or more items per scale, whereas the EXAMINER composite score used 10 items derived from different tasks. Given that sensitivity of the alpha coefficient to the number of items in the scale, it is somewhat impressive that the EXAMINER composite score was able to achieve acceptable reliability as a clinical tool with so few items.

These results contribute to the emerging empirical literature which suggests that the EXAMINER may assess EF reliably across a variety of ages, ethnicities, and disorders (Kramer, 2014; Kramer, et al., 2013). The developers of the EXAMINER utilized IRT methods to measure factors with fewer variables. IRT is focused on individual items; the reliability of a scale is enhanced by containing a small number of non-redundant items that measure a very specific level of the latent variable. Therefore, it would have been more suitable for this study to utilize the IRT method of analysis, thereby comparing similar results with the EXAMINER manual. Future studies may consider using more sophisticated methods of scale development, such as IRT, to investigate the appropriateness of the EXAMINER for use with pediatric SCD. Future researchers should also consider examining test-retest reliability, in children, with other well established measures of EF. These values should then be used to assess cross-cultural examinations for the reliability of the EXAMINER across different cultural groups.

Hypothesis 2A: The strength of association between EXAMINER scores and VIQ from the WJ-III will be similar for African American and White children.

The psychometric literature reveals that there is an association between VIQ and EF measures for pediatric populations (McCarthy, 1972; Ardila et al., 1998; Ardila et al. 2000). Findings suggest that hypothesis 2A was supported in this study. The EXAMINER shows comparable validity across African American youth, compared to established measures of VIQ. Results from this study reveal that the strength of association between the EXAMINER and VIQ on WRAT and WJ-III was similar for African American and White children, when age was held constant. In addition, three of four partial correlation coefficients for the EXAMINER, relative to VIQ on the WRAT and WJ-III were in the direction of higher convergent validity for the African American sample, as compared to the White sample. This suggests that increasing sample size and power would be unlikely to change the outcome of these analyses. Despite the fact that the EXAMINER appeared to demonstrate higher convergent validity for the African American group, there was also greater variability across partial correlation values for the African American sample. However, the similar convergent validity shown between the EXAMINER with the WJ-III and the WRAT between African American and White samples, demonstrates acceptable cultural validity for African American children. These findings are consistent in the multicultural assessment literature. Reynolds & Suzuki (2012) suggest that when psychological test are developed through culturally fair methods, they are more likely to demonstrate relatively equivalent ability levels across ethnic groups and are also more likely correctly identify patterns of performance between different ethnic groups.

Hypothesis 2B: The strength of the association between the validity coefficient for the BRIEF Behavioral Regulation, Metacognition, and Global Executive Composite and the EXAMINER executive composite, working memory factor, cognitive control factor, and fluency factor scores will be similar for African American and White children.

Convergent validity for the EXAMINER was also assessed with the BRIEF. The BRIEF is a cognitive measure that was developed to examine the functional expression of EF through parent or teacher report of day-to-day behavior (Gioia, Isquith, Guy, Kenworthy, 2000). Findings from hypothesis 2B were supported in this study. Overall, the EXAMINER shows similar validity across the African American and White youth, relative to the BRIEF. Results reveal no significant differences for 10 of 12 analyses to compare the strength of the association between the BRIEF and the EXAMINER for the African American and White comparison groups, controlling for age. In general, these findings suggest that partial correlation values were more variable, but in the direction of higher convergent validity for the African American sample, as compared to the White sample. Again, this suggests that increasing sample size and statistical power would be unlikely to reveal a pattern of poorer convergent validity for African-American children. The standard deviation was slightly larger for the African American comparison group on the BRIEF. Differences in variability may have led to slightly lower coefficients for the White sample. However, the increased variability in validity values does not appear to detract from the moderate correlations found in the African American sample, which support validity for this cultural group

It was reasonable to expect that results of hypothesis 2A and 2B would reveal that the EXAMINER demonstrates similar cultural validity to well-established measures, for

African American children. The mission of the EXAMINER test development was to promote methods of measurement that could be useful across diverse ages and ethnicities (Kramer, et al., 2014). Thereby, this study contributes to the emerging empirical literature which suggests that the EXAMINER may have appropriate detection capabilities for EF in the pediatric SCD population (Kramer, 2014; Kramer, et al., 2013). These results are also a noteworthy first-step to understanding how the EXAMINER performs for African Americans children when compared to White youth. Traditional factor analysis allows for insights into the nature of multiple latent variables, however this typically requires larger sample sizes than our study could efficiently recruit (DeVellis, 2003). An important next step is to explore the construct validity of the EXAMINER with confirmatory factor analysis to test whether the data fit the hypothesized measurement model. The sample size in this study was too small to conduct this analysis. Therefore, future investigations should consider concentrating on further enhancing our understanding of cultural and convergent validity for the EXAMINER, by utilizing larger sample sizes of African American children and other cultural groups.

Hypothesis 3A: Children with SCD at high risk for disease-related cognitive effects will show significantly poorer performance on the EXAMINER than demographically matched control children without SCD. The effect size of these differences will be compared to the magnitude of group differences on selected WJ-III.

Overall, findings from this study suggest that Hypotheses 3A were supported. The EXAMINER appears to evidence comparable sensitivity in detecting neurologic decrements in the pediatric SCD population as compared to the WJ-III. Children with SCD revealed poorer performance on the Examiner Battery compared to the African

American comparison group, providing support for hypothesis 3A. Across all EXAMINER scores, statistically significant findings showed that poorer executive performance occurred for children diagnosed with SCD relative to the African American comparison group. Similarly, results revealed statistically significant differences between the SCD group and the demographically matched control group on each of the variables from the WJ-III. Comparison of the magnitude of the observed effects across the EXAMINER and WJ-III W scale measures suggest a slight trend toward larger effects for the EXAMINER, though overall the values were comparable.

Hypothesis 3B: The EXAMINER will show at least as high of sensitivity and specificity in identifying high risk children with sickle-cell disease as the selected measures from the WJ-III relative to the local, demographically matched comparison group.

The ROC sensitivity-specificity analyses showed that the EXAMINER predicted neurologic risk/status at a comparable rate as WJ-III. These findings provide support for hypothesis 3B. Results of the sensitivity-specificity analyses showed statistical significance for the EXAMINER scores and WJ-III variables in the prediction of neurologic status. The areas under the ROC curves for the EXAMINER and WJ-III evidenced moderate accuracy utilizing Swets's (1988) criteria. Both measures demonstrated classification accuracy rates that were slightly lower than 80%. Given that 80% sensitivity and specificity are often used as common standards for accuracy of psychological tools in decision making, the measures fell slightly below this standard. However, the EXAMINER tool at this point in development does not use age-adjusted scores, which would likely improve its ability to differentiate between groups. Likewise,

non-age adjusted scores were used with the WJ-III to provide a more direct comparison between measures. Thus, the EXAMINER appears to have the potential to provide excellent criterion validity, but needs additional work, primarily in the area of collecting of a larger sample for normative purposes.

Taken together, these results highlight the acceptable levels of criterion validity for both the EXAMINER and WJ-III in terms of sensitivity to SCD-related cognitive deficits though for individual-level decision making the EXAMINER is still lacking in age-specific norms. As the WJ-III is a well-established measure for the detection of neurologic deficits with the pediatric SCD population, these findings provide evidence for the EXAMINER's ability to effectively discriminate between children with and without neurologic morbidity. Therefore, this study contributes to the developing literature which supports the use of the EXAMINER with diverse clinical populations, including pediatric SCD (Kramer, 2014; Kramer et al., 2013). In addition, these results further support the assertion that cognitive testing is an efficacious and cost effective method for identifying neurologic deficits in pediatric SCD. Future investigations should focus efforts on further examining the EXAMINER's ability to detect neurologic deficits within pediatric SCD and its use in clinical trials can be justified.

Limitations

Although the EXAMINER appeared to function as well as traditional cognitive measures for each of these analyses, there are some limitations of this study that must be acknowledged. First, the sample sizes were relatively small for the SCD group and African-American controls, which may account for some of the variation in findings regarding internal consistency. Most notably, the magnitude of difference between the

internal consistency of the EXAMINER and WJ-III verbal comprehension and visual matching was significant for African American and White youth. One reason for higher reliability in the WJ-III may the more limited variability of the EXAMINER sample groups compared to the large normative sample for the WJ-III. Data collection for the multi-site study was performed largely by psychologists, focused on collecting clinical samples. It is likely that the neurologic study samples were in some cases collected as samples of convenience, more so than large well-designed normative data sets. Consequently, decreased variability within the study sample may have translated into lower internal consistency observed for the EXAMINER, relative to the WJ-III normative sample.

The Yurdugul, (2008) article suggests that our sample sizes were large enough to adequately produce a sufficiently unbiased estimator of coefficient alpha for the EXAMINER composite score, given the first eigenvalue of greater than four. However, in detecting differences in alpha across groups one must decide what degree of differences is likely to be meaningful. With a large enough sample, it is likely that group differences in reliability could have been shown. Therefore the critical number of participants may not have been large enough to detect internal consistency. Despite the small sample sizes, results suggest that the overall internal consistency and convergent validity for the EXAMINER evidenced acceptable levels across study groups.

Another limitation of this study concerns the lack of detailed demographics provided for children from other sites. Due to investigator error at these sites, the level of parent education was not provided for the children that participated at 10 sites of the multi-site investigation. It was intended that parent education data would be included in

the analyses, as a proxy variable to represent SES. In 2000, the US Census Bureau reported that African Americans experience disproportionate poverty rates, compared to White non-Latino youth. Schatz and colleagues (2004) found results which suggest that children with SCD experience a combination of biopsychosocial risk factors related to the neurologic factors and socioeconomic disadvantage. Failure to include a variable to account for SES limited the ability to examine cultural differences. SES differences may have influenced some of the comparisons of White and African American youth. It is not clear whether the inclusion of measures for SES would have changed the results found in these samples; future investigations should try to obtain information on these variables to improve our understanding of the performance of African American children on the EXAMINER.

Conclusion

As revealed in the literature review, the great variability related to neuropsychological measures used to assess neurologic morbidity has, in part, negatively impacted treatment outcomes for pediatric SCD. The newly developed EXAMINER was developed with the central goal of reliably and validly assessing EF across a wide range of ages and disorders (Kramer et al., 2012). Therefore it was important to evaluate the psychometric properties of the EXAMINER to help determine the suitability of this tool with children diagnosed with SCD relative to an alternate measure. In sum, this study demonstrated the feasibility of using the EXAMINER for children with SCD as a reliable and valid tool in clinical research and intervention trials. In conclusion, findings in this study suggested that the internal consistency, convergent validity, cultural validity, and sensitivity of the EXAMINER demonstrated that it is a useful neuropsychological test for

the pediatric SCD population. It is important to note that as test development continues for the EXAMINER, consistent evaluation of its psychometric properties should be continued with this clinical population; due to an increased risk of subtle problems with learning and cognition, even in the absence of an identifiable insult to the brain.(Schatz et al., 2002).

REFERENCES

Ackerman, B.P. (1987). Attention and memory in children in context-interactive and context-Independent situations. *Journal of Experimental Child Psychology, 44*(2), 198-221.

Adams, R.J., McKie, V.C., Hsu, L., et al. (1998). Prevention of a first stroke by transfusions in children with sickle cell anemia and abnormal results on transcranial Doppler ultrasonography. *N Engl J Med, 339***,** 5–11.

Adams, R.J., Ohene-Frempong, K., Wang, W, (2001). Sickle Cell and the Brain. *Hematology*, 31-45.

Alexander, M., & Stuss, D. (2000). Disorders of frontal lobe functioning. Seminars in *Neurology, 20,* 427–437.

American Education Research Association. American Psychological Association, and the National Council on Measurement in Education *Standards for educational and psychological testing.* American Education Research Association; Washington, D.C.: 1999.

Anderson, P., Anderson, V., & Lajoie, G. (1996). The Tower of London Test: Validation and standardization for pediatric populations. *The Clinical Neuropsychologist, 10,* 54–65.

Anderson, M.C., Green, C. (2001). Suppressing unwanted memories by executive control. *Nature, 410*(6826), 366–9.

Anderson, V. (1998) Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychological Rehabilitation, 8,* 319–349.

Anderson, V.A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. Developmental *Neuropsychology, 20,* 358-406.

Anderson, V. Lajoie, G., Bell, R. (1997). *Neuropsychology assessment of the school age child.* Department of Psychology, University of Melbourne.

Anon. Stroke With Transfusions Changing to Hydroxyurea (SWiTCH). June 25, 2010. http://clinicaltrials.gov/ct2/show/NCT00122980 (accessed Aug 17, 2011).

Archibald, S., Kimberly, A., Kerns, K.A., (1999). Identification and Description of New Tests of Executive Functioning in Children. *Child Neuropsychology, 5*, 2, 115-12Ardila, A., Galeano, L., Rosselli, M. (1998). Toward a model of neuropsychological activity. *Neuropsychology Review, 8*, 171–190.

Ardila, A., Ostrosky-Solis, F., Rosselli, M., Gómez. C. (2000). Age-related cognitive decline during normal aging: the complex effect of education. *Archives of clinical neuropsychology, 15*, 6, 495-513.

Ariffa, S., Lovell, M., Podell, K., & Goldberg, E. (1998). Wisconsin Card Sorting Test performance in above average and superior school children. Relationship to intelligence and age. *Archives of Clinical Psychology, 13,* 713-720.

Armstrong, F. D., Thompson, R. J.,Wang,W., Zimmerman, R., Pegelow, C. H., Miller, S., Moser, F., Bello, J., Hurtig, A., andVass, K. (1996). Cognitive functioning and brain magnetic resonance imaging in children with sickle cell disease. *Pediatrics* 97: 864–870.

Aron AR (June 2007). "The neural basis of inhibition in cognitive control". *Neuroscientist,* 13 (3) 214–28.

Axelrod, B. N., & Henry, R. R, (1992). Age related performance on Wisconsin card sorting, similarities, and controlled oral word association tests. *The Clinical Neuropsychologist, 6*, 16-26.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.) Recent advances in learning and motivation (vol. 8). New York: Academic.

Baldeweg, T., Hogan, A.M., Saunders, D.E., Telfer, P., Gadian, D.G., Vargha-Khadem, F., et al. (2006). Detecting white matter injury in sickle cell disease using voxel-based morphometry. *Ann Neurol, 59,* 662-72.

Ballas, S.K., Carlos, T.M., Dampier, C., et al. *Guidelines for Standard of Care of Acute PainfulEpisodes in Patients with Sickle Cell Disease.* Pennsylvania Department of Health, 2000 Benton, A.L., Hamsher, K.S. (1978). *Multilingual Aphasia Examination*. Iowa City: University of Iowa.

Becker, J. T., Morris, R. G. (1999). "Working Memory". *Brain and Cognition* 41: 1–8.

Berg, C., Edwards, D., & King, A. (2012). Executive function performance on the children's kitchen task assessment with children with sickle cell disease and matched controls. *Child Neuropsychology, 18*(5), 432-448.

Bernaudin, F., Verlhac, S., Freard, F., Roudot-Thoraval, F., Benkerrou, M., Thuret, I., Mardini, R., Vannier, J.P., Ploix, E., Romero, M., Casse-Perrot, C., Helly, M., Gillard, E., Sebag, G., Kchouk, H., Pracros, J.P., Finck, B., Dacher, J.N.,

Ickowicz, V., Raybaud, C., Poncet, M., Lesprit, E., Renert, P.H., & Brugieres, P. (2000). Multicenter prospective study of children with sickle cell disease: Radiographic and psychometric correlation. *Journal of Child Neurology, 15*, 333–343.

Bishop, D. V. M., Aamodt‐Leeper, G., Creswell, C., McGurk, R., Skuse, D. H. (2001). Individual differences in cognitive planning on the Tower of Hanoi task: neuropsychological maturity or measurement error? *Journal of Child Psychology and Psychiatry*, 42, 4, 551 – 556.

Bland J, A., (1997). Statistics notes: Cronbach's alpha. *BMJ*, 315, 275.

Block, J. (1987). Misgivings about the Matching Familiar Figures Test: Premature or Overdue? *Developmental Psychology, 23*, 740-41.

Boni, L.C., Brown, R.T., Davis, P.C., Hsu, L., and Hopkins, K. (2001). Social information processing and magnetic resonance imagine in children with sickle cell disease. *J Pediatr Psychol*, *26*, 309 –319.

Boone, K. B., Gharffarian, S., Lesser, J.M., Hill-Gutierrez, E. & Berman, N. G. (1993). Wisconsin Card Sorting test performance in healthy, older adults; Relationship to age, sex, education, and IQ. *Journal of Clinical Psychology, 49*, 54-60.

Boone, K.B. (1999). Neuropsychological assessment of executive functions: Impact of age, education, gender, intellectual levels, and vascular status on executive test scores. In B. Miller & J.L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (pp 247-260). NewYork: Guilford Press.

Bowman, JE, Murray, RF Jr., 1990. Genetic variation and disorders in peoples of African origin. Baltimore: *Johns Hopkins University Press*, 196-201.

Brandling-Bennett EM, White DA, et al. (2003). Patterns of verbal long-term and working memory performance reveal deficits in strategic processing in children with frontal infarcts related to sickle cell disease. *Dev Neuropsychology, 24*, 423–434.

Brody, N. (1997). Intelligence, schooling, and society. *American Psycholoist, 52*, 1046-1050.

Brown, T.E. (2005). Attention Deficit Disorder: The Unfocused Mind in Children and Adults. New Haven, CT, Yale University Press Health and Wellness. 20 – 58.

Brown, R. T., Armstrong, F. D., and Eckman, J. R. (1993). Neurocognitive aspects of pediatric sickle cell disease. *Journal of Learning Disabilities,* 26, 33–45.

Brown, R. T., Davis, P. C., Lambert, R., Hsu, L., Hopkins, K., and Eckman, J. (2000). Neurocognitive functioning and magnetic resonance imaging in children with sickle cell disease. *Journal of Pediatric Psychology,* 25, 503–513.

Canavan, A.G., Sprengelmeyer, H.C., Homberg, V. (1994). Conditional associative learning is impaired in cerebellar disease in humans. *Behavioral Neuroscience, 108,* 475-485.

Canivez, G.L., Watkins, M.W. (1999). Long-term stability of the Wechsler Intelligence Scale for Children 3rd Edition among demographic subgroups: gender, race/ethnicity, and age. *Journal of Psycho-educational Assessment, 17,* 300-313.

Cash, Dyer, D. (2009). Cultural differences on the children's memory scale. Doctoral dissertation, Texas A&M University. Available electronically from http : / /hdl .handle .net /1969 .1 /ETD -TAMU -1621.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika, 12,* 197-220.

Cattell, J., M. (1888). The psychological laboratory at Leipsic. *Mind, 13,* 37-51.

Cattell, J., M. (1886). The time taken up by cerebral operations. *Mind,* 11, 277-282, 524-538.

Charache, S., Lubin, B., Reid, C.D. (1992). Management and Therapy of Sickle Cell Disease (NIH publication no. 91-2117). Washington, DC: Superintendent of Documents, U.S. Government Printing Office.

Charter, R. A. (1999). Sample Size Requirements for Precise Estimates of Reliability, Generalizability, and Validity Coefficients. *Journal of Clinical and Experimental Neuropsychology, 21,* 559-566.

Chelune, G., & Baer, R. (1986). Developmental norms for the Wisconsin Card Sorting Test. *Journal of Clinical and Experimental Neuropsychology, 8,* 219–228.

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment,* 6, 284-290.

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

Cohen, M. J., Branch, W. B., McKie, V. C., and Adams, R. J. (1994). Neuropsychological impairment in children with sickle cell anemia and cerebrovascular accidents. *Clinical Pediatrics* 33: 517–524.

Cohen, M.J. (1997). *Children's Memory Scale*. San Antonio Texas: The Psychological Corporation.

Conners, C. K. (1994). *The Conners Continuous Performance Test*. Toronto, Canada: Multi-Health Systems.

Cortina, J.M., (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology, 78*(1), 98-104.

Clark, L.A., Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Craft S, Schatz J, Glauser T, et al. (1993). Neuropsychological effects of stroke in children with sickle cell anemia. *J Pediatr, 123*, 712–717.

Craft, S., Schatz, J., Glauser, T., Lee, B., and DeBaun, M. (1994). The effects of bifrontal stroke during childhood on visual attention: Evidence from children with sickle cell anemia *Developmental Neuropsychology, 10*, 285–297.

Crossley, M., M. Hiscock, et al. (2004). "Dual-task performance in early stage dementia: differential effects for automatized and effortful processing." *J Clin Exp Neuropsychol* 26(3): 332-46.

Darlington, R.B. (1971). Another look at "cultural fairness." *Journal of Educational Measurement.* 8, 71–82.

Davies, S.C. (1993). Bone marrow transplantation for sickle cell disease-the dilemma. *Blood Rev 49*, 4-9.

DeBaun, M. R., Derdeyn, C. P. and McKinstry, R. C. (2006), *Etiology of strokes in children with sickle cell anemia*. Ment. Retard. Dev. Disabil. Res. Rev., 12: 192–199.

DeBaun, M. R., Schatz, J., Siegel, M. J., Koby, M., Craft, S., Resar, L., Chu, J.-Y., Launius, G., Dadash-Zadeh, M., Lee, R. B., and Noetzel, M. (1998). Cognitive screening examinations for silent cerebral infarcts in sickle cell disease. *Neurology, 50*, 1678–1682.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1994). *The California Verbal Learning Test-Children's Version*. New York: Psychological Corporation.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan Executive Function System (D-KEFS) EXAMINER's Manual*. San Antonio, TX: Psychological Corporation.

DeVellis R., (2003). *Scale development: theory and applications: theory and application.* Thousand Okas, CA: Sage.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development, 56*, 868–883.

Diamond, A., & Doar, B. (1989). The performance of human infants on a measure of frontal cortex function, the Delayed Response task. *Developmental Psychobiology, 22*, 271–294.

Diamond, A., & Goldman-Rakic, P. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research, 74*, 24–40.

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to ''do as I say, not as I do''. *Developmental Psychobiology, 29*, 315–334.

Diascro, M.N., & Brody, N. (1994). Odd Man Out and Intelligence. *Intelligence, 19*, 79-92.

Diehr MC, Heaton RK, Miller SW, Grant I. (1998). The Paced Auditory Serial Addition Task (PSAT): norms for age, education and ethnicity. *Assessment. 5*, 375-387.

Dikmen, S.S., Heaton, R.K., Grant, I., Temkin, N.R. (2000). Test-retest reliability and practice effects of Expanded Halstead Neuropsychological Test Battery. *Journal of the International Neuropsychological Society. 5*, 346-356.

Diller, L., Ben-Yishay, Y. Gerstman, L.J., Goodkin, R., Gordon, W., & Weinberg, -J. (1974). *Studies in cognition and rehabilitation in hemiplegia (Rehabilitaton Monograph No. 50).* New York: New York University Medical Center Institute of Rehabilitation Medicine.

Donders, F., C. (1969). Over de snelheid van psychische processen [On the speed of psychological processes]. (W. Koster, Trans.), In W. Koster, (Ed.), *Attention and performance:II.* Amsterdam: North-Holland. (Original work published 1868)

Donders, J., & Hoffman, N.M. (2002). Gender differences in learning and memory after pediatric traumatic brain injury. *Neuropsychology, 16*(4), 491-499.

Dupuy, T., & Cenedela, M. (2000). *Test of Variables of Attention: User's Guide:* Los Angeles, California: Universal Attention Disorders.

Duffy, M. L. (1988). A correlational analysis of the Wechsler Intelligence Scale for Children—Revised, the Detroit Test of Learning Aptitude--2, the Wide Range Achievement Test—Revised, and the Gates-MacGinitie Reading Tests in a sample referred for suspected learning disabilities: Dissertation Abstracts International.

Earley, C.J., Kittner, S.J., Feeser, B.R. et al. (1998). Stroke in children and sickle-cell disease: Baltimore–Washington Cooperative Young Stroke Study. *Neurology, 51,* 169–176.

Edwards, C.L., Scales, M.T., Loughlin, C., Bennett, G.G., Harris-Peterson, S., DeCastro, L.M., Whitworth, E., Abrams, M., Feliu, M., Johnson, S., et al. (2005). A brief review of the pathophysiology, associated pain, and psychosocial issues in sickle cell disease. *International Journal of Behavioral Medicine, 12(3),* 171-179.

Epsy, K.A., Kaufman, P.M., McDiarmid, M.D., & Gilsky, M.L. (1999). Executive Functioning in preschool children: Performance on A-not-B and other delayed response format tasks. *Brain and Cognition, 41,* 178-199.

Espino, D.V., Lichtenstein, M.J., Palmer, R.F.m & Hazuda, H.P. (2004). Evaluation of the Mini Mental State Examination's Internal Consistency in a community based sample of Mexican American and European American elders. Results from the San Antonio longitudinal Study of aging. *Journal of the American Geriatric Society, 52,* 822-827.

Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy.* New York: W. W. Norton & Company.

Filskov, S.B., Boll, T.J. (1987). Handbook of Clinical Psychology. New York, John Wiley & Sons.

Flanagan, D. P., & Harrison, P. L. (2005). *Contemporary intellectual assessment: Theories, tests, and issues. (2nd Edition).* New York, NY, The Guildford Press

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Dynda, A. M. (2008). Best practices in cognitive assessment. *Best Practices in School Psychology V,* Bethesda: NASP Publications.

Folstein, M.F., Folstein, S.E., & McHugh, P.R., & Fanjiang, G. (2001). *Mini-Mental State Examination: User's Guide.* Odessa, FL: PAR.

Fowler, M. G., Whitt, J. K., Lallinger, R. R., Nash, K. B., Atkinson, S. S.,Wells, R. J., and McMillan, C. (1988). Neuropsychologic and academic functioning of children with sickle cell anemia. *Developmental and Behavioral Pediatrics,* 9, 213–220.

Franck, L.S., Treadmill, M., Jacob, E., & Vinchinsky, E. (2002). Assessment of sickle cell pain in children and young adults using the adolescent pediatric pain tool. *Journal of Pain and Symptom Management, 23*, 114-120.

Frearson, W., & Eysenck, H.J. (1986). Intelligence, reaction time (RT) and a new "odd man out" RT paradigm. *Personality and Individual Differences, 7*, 807-817.

Gagnadoux, M. F., Loirat, C., Bertheleme, J. P., Maisin, A., Kamoun, A., Dabout, Poisson, D., & Broyer, M. (1994). Treatment of anemia in hemodialyzed children using recombinant human erythropoietin (Eprex). Results of a French multicenter clinical trial. *Nephrologie, 15*, 207–211.

Gaston MH, Verter JI, Woods G, et al. (1986). Prophylaxis with oral penicillin in children with sickle cell anemia. A randomized trial. *N Engl J Med*, 314,1593-9.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.

Gerstadt, C., Hong, Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 2 –7 years old on a Stroop-like day-night test. *Cognition, 53*, 129–153.

Gioia, G., Isquith, P., Guy, S., Kenworthy, L. (2001). Assessment of executive functions in children with neurological impairment. In R. Simeonsson & S. Rosenthal (Eds.) *Psychological and developmental assessment: Children with disabilities and chronic conditions*, 317-356, New York; The Guilford Press.

Gilhooly, K.J., & Logie, R.H. (1980). Age-of-acquistition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods and Instrumentation, 12*, 395-427.

Gladsjo JA, Evans JD, Schuman CC, et al. (1999). Norms for Letter and Category Fluency: demographic corrections for age, education, and ethnicity. *Assessment. 6*, 147-178.

Glauser, T.A., Siegel, M.J., Lee, B.C., DeBaun, M.R. (1995) Accuracy of neurologic examination and history in detecting evidence of MRI-diagnosed cerebral infarctions in children with sickle cell hemoglobinopathy. *J Child Neurol, 10,* 88–92.

Gold, J. I., Johnson, C. B., Treadwell, M. J., Hans, N., & Vichinsky, E. (2008). Detection and Assessment of Stroke in Patients with Sickle Cell Disease: Neuropsychological Functioning and Magnetic Resonance Imaging. *Pediatric Hematology & Oncology*, 25(5), 409-421.

Goonan, B. T., Goonan, L. J., Brown, R. T., Buchanan, I., and Eckman, J. R. (1994). Sustained attention and inhibitory control in children with sickle cell syndrome. *Archives of Clinical Neuropsychology*, 9, 89–104.

Gordon, M., & Mettelman, B. B. (1988). The assessment of attention: I. Standardization and reliability of a behavior-based measure. *Journal of Clinical Psychology, 44*, 682-690.

Gottsdanker, R., & Shragg, G.P. (1985). Verification of Donders' Subtraction Method. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 765-776.

Greene, A.C., Sapp, G.L., Chissom, B. (1990). Validation of the Stanford-Binet Intelligence Scale: Fourth edition with exceptional black male students. *Psychology in the Schools, 27,* 1-35.

Guion, R. M. (1980). On trinitarian doctrines of validity. *Professional Psychology, 11*, 385-398.

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development, 61,* 653–663.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage Press.

Hariman, L.M., Griffith, E.R., Hurtig, A.L., Keehn, M.T. (1991). Functional outcomes of children with sickle-cell disease affected by stroke. *Arch. Phys. Med. Rehabil,. 72*, 498–502.

Hammill, D. D. (1985). *Detroit Tests of Learning Aptitude-2, EXAMINER's Manual*. Austin, TX: Pro-Ed.

Harris, A., Parker, N., & Barker, C. (1998). Adults with sickle cell disease: psychological impact and experience of hospital service. *Psychology, Health, & Medicine, 3*, 171-179.

Hebbel, R. P., Mohandas, N., Embury, S. H., & Hebbel, R. P. (1994). Sickle cell adherence. *Sickle Cell Disease: Basic Principles and Clinical Practice. New York, NY, Raven*.

Helms, J.,E. (2002). A remedy for the Black-White test-score disparity [Comment] *American Psychologist,* 57(4), 303–305.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging*, 18, 639–657.

Huttenlocher, P.R., Moohr, J.W., Johns, L. & Brown, F.D. (1984) Cerebral blood flow in sickle cell cerebrovascular disease. *Pediatrics, 73,* 615–621.

Kail, R. (1986). Sources of age differences in speed of processing. *Child Development, 57,* 969–987.

Kamin, L. J. (1995). The pioneers of IQ testing. In Ressell Jacoby & Naomi Glauberman (Eds.), *The Bell Curve Debate: History, Documents, Opinions.* New York: Times Books.

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

Kaufman, A. S., Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2006). Test Review: Wechsler, D. (2003). "Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV)." San Antonio, TX: The Psychological Corporation. *Journal Of Psychoeducational Assessment, 24*(3), 278-295.

Kennepohl, S., Shore, D., Nabors, N., & Hanks, R. (2004). African American Acculturation and Neuropsychological test Performance following Traumatic Brain Injury. *Journal of International Neuropsychological Society, 10, 566-577.*

King, A, Herron, S, McKinstry, R et al. (2006). A multidisciplinary health care team's efforts to improve educational attainment in children with sickle-cell anemia and cerebral infarcts. *J. Sch. Health 76* 33–37.

King, A.A,, White, D.A., McKinstry, R.C., Noetzel, M., Debaun, M.R. (2007). A pilot randomized education rehabilitation trial is feasible in sickle cell and strokes. *Neurology 68,* 2008–2011.

Klahr, D., & Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cognitive Psychology, 13,* 113-148.

Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. New York: Methune & Company.

Knight, S., Singhal, A., Thomas, P., and Serjeant, G. (1995). Factors associated with lowered intelligence in homozygous sickle cell disease. *Archives of Disease in Childhood 73*: 316–320.

Kral, M. W., Brown, R., Nietert, P.J., Abboud, M.R., Jackson, S.M., Hynd, G.W. 2003. Transcranial Doppler Ultrasonography and Neurocognitive Functioning in Children with Sickle Cell Disease. *Pediatrics, 112*(2), 324.

Kral, M., & Brown, R. (2004). Transcranial Doppler Ultrasonography and Executive Dysfunction in Children with Sickle Cell Disease. *Journal Of Pediatric Psychology*, *29*(3), 185-195.

Kramer, J., Mungas, D., Possin, K., Rankin, K., Boxer, A., Rosen, H., & ... Widmeyer, M. (2014). NIH EXAMINER: conceptualization and development of an executive function battery. Journal Of The International Neuropsychological Society: JINS, 20(1), 11-19.

Krikorian, R., & Bartok, J. (1998). Developmental data for the Porteus Maze Test. *The Clinical Neuropsychologist, 12,* 305–310.

Kojima, H. (1976). Some Psychometric Problems of the Matching Familiar Figures Test. *Perceptual and motor skills, 43,* 3, 731.

Kolen, Michael J., Zeng, Lingjia, Hanson, Bradley A. (1996). Conditional Standard Errors of Measurement for Scale Scores Using IRT. *Journal of Educational Measurement, 33(2),* 129-40.

Kugler, S., Anderson, B., Cross, D., Sharif, Z., Sano, M., Haggerty, R., Prohovnik, I., Hurlet-Jensen, A., Hilal, S., Mohr, J. P., and DeVivo, D. C. (1993). Abnormal cranial magnetic resonance imaging scans in sickle-cell disease: Neurological correlates and clinical implications. Archives of Neurology 50: 629–635.

Lankron S, Stouse JJ, Wilson R, et al. (2008). Systematic review: hydroxyurea for the treatment of adults with sickle cell disease. *Ann Intern Med, 48*, 939–55.

Lee, T.M.C., Cheung, C.C.Y., Chan, L., & Chan, C.C.H. (2000). Trail making across languages. *Journal of Clinical and Experimental Neuropsychology, 22,* 772-778.

Lee, M.T., Piomelli, S., Granger, S., Miller, S.T., Harkness, S., Brambilla, D.J., et al., (2006). Stroke Prevention trial in sickle cell anemia (STOP): extended follow-up and final results. *Blood, 108(3),* 847-852.

Levin, H., Culhane, K., Hartmann, J., Evankovich, K., Mattson, A., Harward, H., Ringholz, G., Ewing-Cobbs, L., & Fletcher, J. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology, 7,* 377–395.

Lezak, M. D. (1983). Neuropsychological assessment (2nd ed.). New York: Oxford University Press.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J. & Fischer, J. S. (2004).*Neuropsychological Assessment, 4th ed.*. Oxford: Oxford University Press.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. Psychological Reports, 3, 634-694.

Lowry, Richard. "VassarStats: Statistical Computation Web Site." *VassarStats: Statistical Computation Web Site*. Vassar College, 1998. Web. 05 May 2014.

Lucas, J.A., Ivnik, R.J., Smith, G.E., Ferman, T. J., Willis, F.B., Petersen, R.C., & Graft-Radford, N.R. (2005). Mayo's Older African Americans Normative Studies: Norms for Boston Naming Test, Controlled Oral Word Association, Category Fluency, Animal Naming, Token Test, WRAT-3 Reading, Trail Making Test, Stroop Test, and Judgment of Line Orientation. *The Clinical Neuropsychologist, 19*, 243-269.

Massaro, D.,W. (1975). Experimental Psychology and Information Processing, Rand McNally.

Matarazzo J. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *Am Psychol. 45*, 999–1017.

Manly, J.L., Miller, S.W., Heaton, R.K., Byrd, D. Reilly, J., Velasquez, R.L., Saccuzzo, D.P., Grant, L., and the HIV Neurobehavioral Research Center (HNRC) Group. (1998). The effect of African American acculturation on neuropsychological test performance in normal and HIV-positive individuals. *Journal of International Neuropsychological Society, 4*, 291-302.

McCrae, J., D., & Lumley, M. A. (1998). Health status in sickle cell disease: examining the roles of pain coping strategies, somatic awareness, and negative affectivity. *Journal of Behavioral Medicine, 21*, 35-55.

McCarney D, Greenberg LM (1990), Test of Variables of Attention. Minneapolis: Attention Tech Inc. McGrew, K.S., 7 Woodcock, R.W. (2001). *Woodcock Johnson III Technical Manual.* Itasca, IL: Riverside Publishing.

McClure, F. D., & Gordon, M. (1984). The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. *Journal of Abnormal Child Psychology*, 12, 561-572.

McGrew, K., & Flanagan, D. (1998). *The Intelligence Test Desk Reference (ITDR). Gf-Gc cross-battery assessment*. Boston. Allyn & Bacon.

McIntire, S. A. & Miller, L. A. (2007). *Foundations of psychological testing: a practical approach. 2nd ed.* Thousand Oaks, CA: Sage Publications.

Meeker, William Q. and Luis A. Escobar. (1998). *Statistical Methods for Reliability Data*. Hoboken, New Jersey: Wiley.

Mehta, S.R., Afenyi-Annan, A., Byrns, P.J., Lottenberg, R. (2006). Opportunities to improve outcomes in sickle cell disease. *Am Fam Physician. 74,* 2, 303-10.

Meyer G, Finn S, Eyde L, et al. (2001). Psychological testing and psychological assessment: a review of evidence and issues. *Amer Psychol, 56,* 128–165.

Miller, S.T., Sleeper, L.A., Pegelow, C.H. Enos, L.E., Wang, W.C., Weiner, S.J. Wethers, D.L., Smith, J, and Kinney, T.R. (2000). Prediction of Adverse Outcomes in Children with Sickle Cell Disease. *N Engl J Med. 342,* 83-89.

Mohandas N, Hebbel R. Pathogenesis of hemolytic anemia. In: Embury SH, et al., eds. 1994 *Sickle cell disease: basic principles and clinical practice*. New York: Raven, 1994:327- 34.

Moscovitch M., (1992). Memory and working with memory: A component process model based on modules and central systems. *J Cogn Neurosci, 4,* 257–267.

Mumford, M.D., Marks, M.A., Connelly, M.S., Zaccaro, S.J., Reiter-palmon, R. (2000). Development of leadership skills: Experience and timing. *The Leadership Quarterly, 11*(1), 87-114.

Nettles, A. L. (1994). Scholastic performance of children with sickle cell disease. *Journal of Health and Social Policy, 5*(3–4), 123–140.

Nikhar, H.S., Meshram, S.U., Shinde, G.B. (2011). A comprehensive review of diverse issues related to sickle cell disease. *East Afr J Public Health, 8*(2), 164-9.

Noll, R. B., Stith, L., Gartstein, M. A., Ris, M. D., Grueneich, R., Vannatta, K., & Kalinyak, K. (2001). Neuropsychological function of youths with sickle cell disease: Comparison with non-chronically ill peers. *Journal of Pediatric Psychology, 26*(2), 69-78.

Norman, D. A. and Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. *Consciousness and self-regulation: Advances in research and theory*. R. J. D. e. al. New York, Plenum Press.

Norman M.C., Evans J.D., Miller S.W., Heaton R.K.. Demographically corrected norms for the California Verbal Learning Test. Journal of Clinical and Experimental *Neuropsychology. 2000,* 22, 80-94.

Nunnally, J.C.&Bernstein, I.H. (1994). *Psychometric theory (3rd ed.).* New York: McGraw-Hill.

O'Brien A, Chiaravalloti N, Goverover Y, DeLuca J. Evidenced based cognitive rehabilitation for persons with multiple sclerosis: a review of the literature. Arch Phys Med Rehahil 2008; 89: 761-69.

Ohene-Frempong K: (1991). Stroke in sickle cell disease: Demographic, clinical and therapeutic considerations. Semin Hematol, 28, 213.

Ouvrier, R.A., Goldsmith, R.F., Ovrier, S., & Williams, L.C. (1993). The value of the Mini-Mental State Examination in childhood: A preliminary study. *Journal of Child Neurology, 6,* 145-149.

Palermo, T.M., Schwartz, L., Drotar D. and McGowan, K. Parental Report of Health Related Quality of Life in Children with Sickle Cell Disease. *Journal of Behavioral Medicine, 25,* 3, 269-283.

Panepinto, J.A., Walters, M.C., Carreras, J., Marsh, J., et al. (2007). Matched-related donor transplantation for sickle cell disease: report from the Center for International Blood and Transplant Research. *British Journal of Haematology, 137,* 5, 479-485.

Pegelow, C..H, Adams, R.J., McKie, V., et al. (1995). Risk of recurrent stroke in patients with sickle cell disease treated with erythrocyte transfusions. *J Pediatr. 126,* 896-899.

Platt, O. S., Brambilla, D. J., Rosse, W. F., Milner, P. F., Castro, O., Steinberg, M. H., et al. (1994). Mortality in sickle cell disease. Life expectancy and risk factors for early death. *New England Journal of Medicine, 330,* 23, 1639–1644.

Peterson, R.A. (1994) A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21,* 381-391.

Petrides, M., & Milner, B. (1982). Deficits on subject ordered tasks after frontal and temporal-lobe lesions in man. Neuropsychologia, 20, 249-262.

Piomelli, S. Hart, D. Graziano, J., Grant, G., Karpatkin, , M., McCarthy, K. (1985). Current Strategies in the Management of Cooley's Anemia. *Annals of the New York Academy of Sciences, 445,* 256.

Pomati, S., Farina, E., Magni, E., Laiacona, M., Mariani, C. (1996). Normative data for two neuropsychological tests sensitive to frontal dysfunction. *Italian Journal of Neurological Sciences, 17,* 201-209.

Posner, M.I. (1980). Orienting of attention. Quarterly *Journal of Experimental Psychology, 41A,* 19-45.

Posner, M.I., Snyder, C.R.(1975). "Attention and cognitive control". In Solso RL. *Information processing and cognition: the Loyola symposium.* Hillsdale, N.J: L. Erlbaum Associates.
.

Posner, M.I., Petersen, S.E. (1990). "The attention system of the human brain". *Annu Rev Neurosci*, *13*(1), 25–42.

Puffer, E.S., Schatz, J.C., Roberts, C.W. (2007). The Association of Oral Hydroxyurea Therapy with Improved Cognitive Functioning in Sickle Cell Disease. *Child Neuropsychology, 13,* 142-154.

Puffer, E.S., Schatz, J.C., Roberts, C.W. (2010). Relationships between Somatic Growth and Cognitive Functioning in Young Children with Sickle Cell Disease. *Journal of Pediatric Psychology*, *35*, 892–904.

Raake, A. (2006) Front Matter, in Speech Quality of VoIP: Assessment and Prediction, John Wiley & Sons, Ltd, Chichester, UK.

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1-38). UK: Psychology Press.

Raj, A., Bertolone, S. J., Mangold, S., Edmonds, H. L. Jr. (2004). Assessment of cerebral tissue Oxygenation in patients with sickle cell disease: Effect of transfusion therapy. Journal of *Pediatric Hematology Oncology, 26,* 279–283.

Reitan, R.M., & Wolfson, D. (1985). The Halstead-Reitan Neuropsychological Test Battery. Tucson, Arizona: Neuropsychological Press.

Rey, G.J., Feldman, E., Rivas-Vasquez, R., Levin, B.E., & Benton, A. (1999). Neuropsychological test development and normative development and normative data on Hispanics. *Archives of Neurology*, *59,* 639-647.

Reynolds, Cecil R.; Goldstein, Sam (1999). *Handbook of neurodevelopmental and genetic disorders in children*. New York: The Guilford Press. pp. 3–8.

Riccio, C.A., Reynolsa, C.R. & Lowe, P.A. (2001). *Clinical applications of continuous performance tests:Measuring attention and impulsive reading in children and adults.* New York: John Wiley & Sons.

Russell MO, Goldberg HI, Reis L, et al. (1976). Transfusion therapy for cerebrovascular abnormalities in sickle cell disease. *J Pediatr. 88,* 382-387

Robertson, L. C., Egly, R., Lamb, M. R., & Kerth, L. (1993). Spatial attention and cuing to global and local levels of hierarchical structure. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 471–487.

Rocke, K., Hays, P., Edwards, D., & Berg, C. (2008). Development of a performance assessment of executive function: the children's kitchen task assessment. *AJOT: American Journal Of Occupational Therapy*, (5), 528.

Rönnlund, M., Lövdén, M., & Nilsson, L. G. (2001). Adult age differences in Tower of Hanoi Performance: Influence from demographic and cognitive variables. Aging, *Neuropsychology & Cognition, 8,* 269-283.

Ruff, R. M., C. C. Allen, et al. (1994). "Figural fluency: differential impairment in patients with left versus right frontal lobe lesions." *Arch Clin Neuropsychol* 9(1): 41-55.

Salkind, N. J., Kojima, H., and Zelniker, T, (1978). Cognitive Tempo in American, Japanese, and Israeli Children. *Child Development, 49, 4,* 1024-1027.

Seeler, R.A., Royal, J.E. (1977). Acute and chronic management of children with sickle cell anemia and cerebrovascular occlusive crisis. *Ill Med J. 151,* 267-269

Schatz, J. (2004). Brief report: Academic attainment in children with sickle cell disease. *Journal of Pediatric Psychology*, *29*, 627– 633.

Schatz, J., Brown, R.T., Pascual, J.M., et al. (2001). Poor school and cognitive f unctioning with Silent cerebral infarction and sickle cell disease. *Neurology, 56,* 1109 –1111.

Schatz, J., Buzan, R.F. (2006). Decreased corpus callosum size in sickle cell disease: Relationship with cerebral infarcts and cognitive functioning. *J Int Neuropsycho Soc, 12*, 24 –34.

Schatz, J., Craft, S., Koby, M., et al. (1999). Neuropsychological deficits in children with sickle cell disease and cerebral infarction: The role of lesion location and volume. *Child Neuropsychology, 5*, 92–103.

Schatz, J., Craft, S., Koby, M, et al. (2004). Asymmetries in visual-spatial processing following childhood stroke. *Neuropsychology, 18*, 340–352.

Schatz, J., Finke, R.L., Kellett, J.M., & Kramer, J.H. (2002). Cognitive functioning in children with sickle cell disease: A meta-analysis. *Journal of Pediatric Psychology, 8*, 739–748.

Schatz, J., Finke, R.L., Roberts, C.W. (2004). Interactions among biomedical and environmental factors in cognitive development: A preliminary study of sickle cell disease. *Journal of Developmental and Behavioral Pediatrics, 25*, 303–310.

Schatz J, Koby M, Siegel MJ, et al. (1999). Neuropsychologic deficits in children with sickle cell disease and cerebral infarction: Role of lesion site and volume. *Child Neuropsychol, 5*, 92-103.

Schatz J, Roberts CW. (2005). Short-term memory in children with sickle cell disease: Executive versus modality specific processing deficits. *Arch Clin Neuropsychol, 20,* 1073–1085.

Schatz J, Roberts, C.W. (2007). Neurobehavioral impact of sickle cell disease in early childhood. *J Int Neuropsychol Soc, 13(6):*933-43.

Schatz J, White DA, Moinuddin A, et al. (2002). Lesion burden and cognitive morbidity in children with sickle cell disease. *J Child Neurol 17*, 891– 895.

Schmidt, M. (1996). *Rey Auditory-Verbal Learning Test.* Los Angeles: Western Psychological Services.

Schum, R.L., Sivan, A.B., & Benton, A. (1989). Multilingual Aphasia Examination: Norms for Children. *Clinical Neuropsychologist*, 3, 375-383.

Segall, D. O. (1994). The reliability of linearly equated tests. *Psychometrika, 59,* 361-375.

Serjeant, G. R. (1997). Sickle-cell disease. *Lancet, 350,* 725-730.

Shapiro, B.K., Batshaw, M.L.. (2011). Intellectual disability. In: Kliegman RM, Behrman RE, Jenson HB, Stanton BF, eds. *Nelson Textbook of Pediatrics* . 19th ed. Philadelphia, Pa: Saunders Elsevier; chap 33.

Shiffrin, R.M., Schneider, W. (1977). "Controlled and automatic human information processing: II: Perceptual learning, automatic attending, and a general theory". *Psychological Review,* 84(2), 127–90.

Shu, B.C., Tien, A.Y., Lung,F.W., & Change, Y.Y. (2000). Norms for the Wisconsin Card Sorting Test in 6 to 11 year old children in Taiwan. *The Clinical Psychologist, 14,* 275-286.

Siegler, R. S. (1992). The other Alfred Binet. *Developmental Psychology, 28,* 179-190.

Sickle Cell Disease Guideline Panel. 1993, Sickle cell disease: screening, diagnosis, management, and counseling in newborns and infants. Rockville, Md.: U.S. Dept. of Health and Human Services, Agency for Health Care Policy and Research; *AHCPR publication no. 93*-0562.

Sparrow, S.S., Davis, S.M. (2000). Recent Advances in the Assessment of Intelligence and Cognition. *J. Child Psychol. Psychiat, 41*, 117-131.

Strauss, E., Sherman, E.M.S., Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. Oxford University Press.

Steen, R.G., Fineberg-Buchner, C., Hankins, G., Weiss, L., Prifitera, A., & Mulhern, R.K. (2005). Cognitive deficits in children with sickle cell disease. *Journal of Child Neurology, 20,* 102–107.

Steen, R.G., Hu, X.J., Elliott, V.E., et al. (2002). Kindergarten readiness skills in children with Sickle cell disease: Evidence of early neurocognitive damage? *J Child Neurol, 17,* 111–116.

Steen R.G., Miles, M.A., Helton, K.J., et al. (2003). Cognitive impairment in children with Hemoglobin SS sickle cell disease: Relationship to MR imaging findings and hematocrit.*Am J Neuroradiol, 24,* 382–389.

Steen, R. G., Reddick,W. E., Mulhern, R. K., Langston, J.W., Ogg, R. J., Bieberich, A. A., Kingsley, P. B., and Wang,W. C. (1998). Quantitative MRI of the brain in children with sickle cell disease reveals

Steen, R. G., Xiong, X., Mulhern, R. K., Langston, J. W., and Wang, W. C. (1999). Subtle brain abnormalities in children with sickle cell disease: Relationship to blood hematocrit. *Annals of Neurology, 45,* 279–286.

Sternberg, S. (1969). "The discovery of processing stages: Extensions of Donders'method". *Acta Psychologica,* 30, 276–315.

Stevens, M. C., Maude, G. H., Cupidore, L., Jackson, H., Hayes, R. J., & Serjeant, G. R. (1986). Prepubertal growth and skeletal maturation in children with sickle cell disease. *Pediatrics, 78,* 124–132.

Stewart, L. H., & Kaminski, R. (2002). Best practices in developing local norms for academic problem solving. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology ta IV* (pp. 737–752). Bethesda, MD: National Association of School Psychologists.

Stone, V.E., Gerrans, P. (2006). What's domain-specific about theory of mind? *Soc Neurosci,*1, (4), 309–19.

Strouse JJ, Lanzkron S, Beach MC, et al. (2008). Hydroxyurea for sickle cell disease: a systematic review for efficacy and toxicity in children. *Pediatrics, 122,* 1332–42.

Strub, R.L. and Black, F.W. (1977). The Mental Status Exam in Neurology, Davis, Philadelphia.

Sue, S., Akutsu, P. D., & Higashi, C. (1985). Training issues in conducting therapy with ethnic-minority clients. In P. B. Pedersen (Ed.), *Handbook in cross-cultural counseling and therapy* (pp. 275-280). Westport, CT: Greenwood Press.

Reynolds, C.R., & Suzuki, L.A. (2012). Bias in Psychological Assessment. In I. B. Weiner, J. R. Graham, J. A. Naglieri, *Handbook of Psychology, Assessment Psychology* (pp. 82-108). Hoboken, NJL John Wiley & Sons, Inc.

Swets JA. (1988). Measuring the accuracy of diagnostic systems. *Science, 240*, 1285-93.

Swift, A. V., Cohen, M. J., Hynd, G. W., Wisenbaker, J. M., McKie, K. M., Makari, G., and McKie, V. C. (1989). Neuropsychologic impairment in children with sickle cell anemia. *Pediatrics* 84: 1077–1085.

Talley, J.L., (1990). *Children's auditory verbal learning test: Professional Manual*. Psychological Assessment Resources. Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). *Stanford-Binet Intelligence Scale (4th ed.)*. Chicago: Riverside.

Tarazi, R.A., Grant, M.L., Ely, E., & Barakat, L.P. (2007). Neuropsycholgical functioning in preschool age children with sickle cell disease: the role of illness related and psychosocial factors. *Child Neuropsychology, 13*, 155-172.

Terrell, F., Terrell, S. L. and Taylor, J. (1980), Effects of race of EXAMINER and type of
reinforcement on the intelligence test performance of lower-class black children. *Psychol. Schs., 17*: 270–272

Thompson, R.J., Jr., Gustafson, K.E., Bonner, M.J., & Ware, R.E. (2002). Neurocognitive development of young children with sickle cell disease through three years of age. *Journal of PediatricPsychology, 27*, 235–244.

Tipper ,S.,P. (2001). Does negative priming reflect inhibitory mechanisms? A review and integration of conflicting views. *Q J Exp Psychol A*, 54(2), 321–43.

Ulrich, R., Mattes, S., & Miller, J. (1999). Donders_s assumption of pure insertion: An evaluation on the basis of response dynamics. *Acta Psychologica, 102*, 43–75.

Van der Elst, W., van Boxtel, p.J., van Breukelen, J.P. & Jolles, E. (2005). Rey's Verbal Learning Test: Normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society, 11*, 290-302.

Van der Linden WJ, Hambleton RK. Item response theory: Brief history, common models, and extensions. In: van der Linden WJ, Hambleton RK, editors. *Handbook of modern item response theory.* Springer-Verlag; New York: 1997. pp. 1–28.

Varni JW, Jacobs JR, Seid M. (2000). Treatment adherence as predictor of health-related quality of life. In: Drotar D (ed.) *Promoting Adherence to Medical Treatment in Chronic Childhood Illness: Concepts, methods and interventions.* Mahwah, NJ: Lawrence Erlbaum: 287–305.

Vichinsky, E.P., Neumayr, L.D., Gold, J.I., Weiner, M.W., Rule, R.R., Truran, D., Kasten, J., Eggleston, B., Kesler, K., McMahon, L., Orringer, E.P., Harrington, T.,. Kalinyak, K., De Castro, L.M., Kutlar, A., Rutherford, C.J., Johnson, C., Bessman, J.D., Jordan, L.B., Armstrong, F.D. (2010). Neuropsychological dysfunction and neuroimaging abnormalities in neurologically intact adults with sickle cell anemia. *JAMA, 303,* 1823-31.

Walters MC, Patience M, Leisenring W, Rogers ZR, Dinndorf P, Davies SC, Roberts IAG, Yeager, A, Kurtzberg J, Bunin N, Scott JP, Wall DA, Wayne AS, Wiley J, Darbyshire PJ, Mentzer WC, Smith FO, and Sullivan KM. (1997). Collaborative multicenter investigation of marrow transplantation for sickle cell disease: current results and future directions. *Biol Blood Marrow Transplant, 3,* 310-315, 1997.

Wang, W., Enos, L., Gallagher, D., Thompson, R., Guarini, L., Vichinsky, E., Wright, E., Zimmerman, R., & Armstrong, F.D. (2001). Neuropsychologic performance in school-aged children with sickle cell disease: A report from the Cooperative Study of Sickle Cell Disease. *Journal of Pediatrics, 139,* 391–397.

Wang, W. C., Langston, J. W., Steen, R. G., Wynn, L. W., Mulhern, R. K., Wilimas, J. A., Kim, F. M., and Figueroa, R. E. (1998). Abnormalities of the central nervous system in very young children with sickle cell anemia. *Journal of Pediatrics, 132,* 994–998.

Ware, R.E., Zimmerman, S.A., Schultz, W.H. (1999). Hydroxyurea as an alternative to blood transfusions for the prevention of recurrent stroke in children with sickle cell disease. *Blood, 94,* 3022-3026

Wasserman, A. L., Wilimas, J. A., Fairclough, D. L., Mulhern, R. K., andWang,W. (1991). Subtle neuropsychological deficits in children with sickle cell disease. *American Journal of Pediatric Hematology/Oncology*, 13, 14–20.

Watkins, K. E., Hewes, D. K. M., Connelly, A., Kendall, B. E., Kingsley, D. P. E., Evans. J. E. P., Gadian, D. G., Vargha-Khadem, F., and Kirkham, F. J. (1998). Cognitive deficits associated with frontallobe infarction in children with sickle cell disease. *Developmental Medicine and Child Neurology*, 40: 536–543.

Weatherall D, Hofman K, Rodgers G, Ruffi n J, Hrynkow S. A case for developing North-South Partnerships for research in sickle cell disease. *Blood* 2005; **105**: 921–23.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology, 19,* 87-95.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI) manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale – third edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children–Fourth Edition*. San Antonio, TX: Pearson Education, Inc.

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

Welsh, M.C., Revilla, V., Strongin, D., Kepler, M. (2000). Towers of Hanoi and London: is the nonshared variance due to differences in task administration? *Percept Mot Skills, 90,* 2, 562-72.

White, D. A., and DeBaun, M. (1998). Cognitive and behavioral function in children with sickle cell disease: A review and discussion of methodological issues. *Journal of Pediatric Hematology/Oncology, 20,* 458–462.

White, D. A., Moinuddin, A., McKinstry, R.C., Noetzel, M., Armstrong, M., DeBaun, M. (2006). Cognitive screening for silent cerebral infarction in children with sickle cell disease. *Journal of pediatric hematology/oncology, 28,* 166.

White, D. A., Salorio, C. F., Schatz, J., DeBaun, M. (2000). Preliminary study of working memory in children with stroke related to sickle cell disease. *Journal of clinical and experimental neuropsychology, 22,* 257.

Wilson, V.L., Nolan, R. F., Reynolds, C. R., Kamphaus, R.W. (1989). Race and gender effects on item functioning on the Kaufman assessment battery for children. *Journal of School Psychology, 27,* 3, 289-296.

Woodcock R.W., Johnson M.B. (1989). WJ-R Tests of Cognitive Ability. Itasca, IL: Riverside Publishers.

Woodcock, R. W., & Johnson, M. B. (1977). Woodcock-Johnson Psychoeducation Battery. Allen, TX: DLM Teaching Resources.

Woodruff-Pak, D.D. (1997). *The neuropsychology of aging.* Oxford: Blackwell.

Yerys, B.E., White, D.A, Salorio, C.F., McKinstry, R., Moinuddin, A., DeBaun, M. (2003). Memory strategy training in children with cerebral infarcts related to sickle cell disease. *J. Pediatr. Hematol. Oncol. 25,* 495–498.

Yurdugul, H. (2008). Minimum Sample Size for Cronbach's Coefficient Alpha: A Monte-Carlo Study. *Hacettepe Üniversitesi Journal of Education, 35,* 397-405.

Zelazo, P. D., & Muller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.) *Blackwell handbook of childhood cognitive development*, 445–469, Malden, MA: Blackwell.

Zemel, B. S., Kawchak, D. A., Fung, E. B., & Ohene-Frempong, K. (2002). Effect of zinc supplementation on growth and body composition in children with sickle cell disease. *American Journal of Clinical Nutrition, 75*, 300–307.