2009

An Investigation of the Relationship of Student Performance on the Wisconsin Card Sorting Test with Other Clinical Measures

Lisa A. Perkins
Philadelphia College of Osteopathic Medicine, lperkins6@aol.com

Philadelphia College of Osteopathic Medicine

Department of Psychology

AN INVESTIGATION OF THE RELATIONSHIP OF STUDENT PERFORMANCE ON THE WISCONSIN CARD SORTING TEST WITH OTHER CLINICAL MEASURES

By Lisa A. Perkins

Submitted in Partial Fulfillment of the Requirements of the Degree of

Doctor of Psychology

August 2009

PHILADELPHIA COLLEGE OF OSTEOPATHIC MEDICINE
DEPARTMENT OF PSYCHOLOGY

Dissertation Approval

This is to certify that the thesis presented to us by Lisa Perkins on the 29th day of January, 2009, in partial fulfillment of the requirements for the degree of Doctor of Psychology, has been examined and is acceptable in both scholarship and literary quality.

Committee Members' Signatures:

George McCloskey, Ph.D., Chairperson

James Brad Hale, Ph.D.

Dr. Rocco Marotta

Robert A. DiTomasso, Ph.D., ABPP, Chair, Department of Psychology

Acknowledgements

I would like to acknowledge my doctoral committee: Dr. George McCloskey, Dr. Brad Hale, and Dr. Rocco Marotta. It was my hope that each of these individuals would provide—in different forms—support, guidance, encouragement, and direction in this process and these hopes were certainly fulfilled. Special gratitude goes to George McCloskey, whose frequent cajoling “You are almost there” served to maintain optimism as well as improve performance. I also thank George, Mary Bernhardt, and Bill Blank for the provision of archival data which made this study possible.

I would also like to express gratitude for the support and encouragement provided by friends and family. I especially thank my husband, Andy, who was eternally patient about the hours of work that the process required. Great appreciation goes to the “Wilder Women” and to my Westport and Family Study Center colleagues, whose enthusiasm, expressed pride, and joy at the prospect of my completion of this process was a considerable impetus to carry on.

Abstract

The Wisconsin Card Sorting Test (WCST) is widely used for neuropsychological assessment of executive functions. Although the literature notes that the WCST is a measure of abstract reasoning and cognitive flexibility, there has been little data relative to the constructs that are assessed when the test is used with children or to the relationship between WCST performance and performance on other child assessment tools. This study of 94 children and adolescents referred for psychological evaluations investigated the relationship between scores obtained on the WCST and scores from child and adult versions of the Wechsler intelligence scales and the Delis-Kaplan Executive Function System, and selected Scale scores from the Behavior Rating Inventory of Executive Functions and the Behavior Assessment System for Children, based on ratings obtained from parents and teachers. The results suggest a modest relationship between WCST scores and the scores from other cognitive tests and rating scales, although students generally received higher Wechsler Matrix Reasoning scores than WCST Conceptual Level Response scores. In reviewing the performance of these 94 students, results also suggest differences in performance, based on test administration format (computer or manual) and by age groupings, with older students outperforming younger in spite of age-corrected scores. Analysis of response patterns, particularly the number of trials to complete the second set, and of performance observations suggest that use of a process approach may be helpful in identifying set-shifting and sustained attention difficulties of students not otherwise identified, using the existing WCST scoring procedures. This study suggests that the WCST may offer unique and important insights into the executive function capacities of children and adolescents.

Table of Contents

List of Tables

Chapter One

Introduction

Within the past decade, there has been an increasing focus on the difficulties that children and adolescents have in their academic and daily functioning that appear to be caused by deficits in executive function capacities. The understanding and definition of what constitutes executive functions have undergone considerable study and refinement during this time. Early discussions centered on the construct as a unitary system (Fischer & Daley, 2007). Increasingly, executive functions were thought to be a collection of control processes responsible for goal-directed behavior and problem solving (Gioia, Isquith, & Guy, 2001); simply, they were thought of as ―control processes,” primarily characterized by inhibition and ability to delay response (Denckla, 1996), or they were believed to be coordination of metacognitive and cognitive processes involved in task analysis and strategy control and monitoring (Borkowski & Burke, 1996). More current discussions still offer an overall global definition, but also offer increasing specificity relative to the individual and potentially discrete processes or functions. There is increasing consensus that the global construct of executive functions is a ―collection of interrelated functions that are responsible for purposeful, goal-directed, problem-solving behavior” (Gioia et al., 2001, p. 320). With the increasing specificity of the multiple aspects of executive functions, there also is an increasing awareness that executive function capacities impact not only the ability to complete tasks in the work or educational realm, but also impact the ability to navigate the various dimensions of one‘s environment, be these physical, interpersonal, or intrapersonal (McCloskey, Perkins, & Van Divner, 2009).

Research has long implicated frontal lobe damage or dysfunction with difficulties in aspects of behavioral control, and more specifically, in goal-directed behaviors. Although the brain basis of behavior has been a topic of research for more than a century, the advances in technology over the past few decades have brought increasing specificity to the discussion of the brain/behavior relationships. In his early discussions of the human frontal lobes, Luria (1973) organized their functions into three basic categories: the regulation of states of activity, the regulation of conscious movements and actions, and the regulation of working memory and problem-solving activities. More recent discussions suggest that the frontal lobes, specifically the prefrontal cortex (PFC), support the cognitive functions that coordinate the execution of action, with an emphasis on the coordination of executive processes.

Kaplan, Sengor, Gurvit, Genc, and Guzelis (2006) offer a rather elegant description of the responsibility of the prefrontal cortex, in noting “the PFC has access to a wide variety of refined information about the external physical world and the internal milieu of the organism and holds a unique position for orchestrating often conflicting demands of external reality and internal drives which are essential for voluntary goal-directed behavior” (p. 376). Luria (1973) and others highlight the way in which the physical structure of the prefrontal region has the strongest connections with other regions of the brain, receiving stimuli from the brain stem, the hypothalamus, the limbic system, the thalamus, the basal ganglia, and other subcortical areas. Gioia et al. (2001)

note the intentional use of the term frontal “system” rather than frontal lobes in their discussion of brain/executive functioning to reflect more accurately the interconnections with these other areas. A discussion of the frontal lobes and their connectivity to other brain structures reinforces both the “executive” and “functions” nature of the construct, because this frontal system engages or directs all the components that are required in volitional activity (Hale & Fiorello, 2004). Given this complex interaction process, caution should be exercised in viewing the term frontal or prefrontal lobe functioning as being synonymous with executive functions (Denckla, 1996); however, the literature is consistent in linking the structure and functions. In his *Developmental Variations and Learning Disorders* (1999), Levine quoted Tranel, Anderson, and Benton in saying, “There is rarely a discussion of disturbances of executive functions that does not make reference to dysfunction of prefrontal brain regions” (p. 44).

With this increased understanding of what constitutes the brain basis of executive function difficulties, there has also been greater attention to and incremental development of standardized assessments and ecological measures of these executive function capacities. Within the last decade, standardized instruments, such as the NEPSY and NEPSY-2 (Korkman, Kirk, & Kemp, 1998; 2007), the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001), and indirect measures, such as the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000), have been developed. Of all these measures, the Wisconsin Card Sorting Test (WCST), developed in 1948 as an assessment of abstract reasoning and cognitive flexibility (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), is considered the “gold standard” for neuropsychological assessment of executive functions (Delis et al.). Although these varied assessment instruments are often composed of different tasks or tests, their authors suggest that these tests are assessing similar executive function capacities. Measures of shifting, inhibiting, sustaining, self-monitoring, planning, hypothesis testing, and initiating are found on subtests of the Delis-Kaplan Executive Function System (D-KEFS); planning, inhibiting, sustaining, initiating, shifting, and working memory, on the NEPSY Attention/Executive domain; and working memory, shifting, inhibiting, and sustaining on the Wisconsin Card Sorting Test (WCST). The BRIEF provides scores around eight executive skills: inhibition, emotional control, shifting, initiation, working memory, planning/organization, organization of materials, and self-monitoring. Although the Behavior Assessment System for Children (BASC, BASC-2: Reynolds & Kamphaus, 1993, 2004) is considered a more general clinical and adaptive rating scale, certain items address difficulties with self-regulation of behavior, task planning and organization, and ability to adapt to changing conditions. Although these instruments appear to be measuring similar constructs, few studies have been conducted that report correlations among these measures within the same groups of individuals.

Given the similarity of the descriptions of what is being measured by these assessment tools, it would seem likely that a similar profile of executive function strengths and/or weaknesses would emerge from the scores obtained with all of these assessments; e.g., that a lack of inhibition noted in referral concerns and/or behavioral ratings would also be reflected in the scores earned on each of these instruments.

However, this certainly is not always the case. As noted by Gioia et al.: –A paradox in the assessment of executive functions is that some individuals with significant deficits in specific executive function subdomains may, in fact, perform appropriately on many purported ‗tests of executive function,‘ yet have significant problems making simple real-life decisions" (2000, p. 338) Some clinicians and researchers (e.g., Goldberg & Podell, 2000) posit that the lack of ecological validity, or at least the lack of consistency between test scores and behavior is due to the fact that most neuropsychological instruments, including the WCST, are structured for eliciting specific predetermined correct responses; such veridical decision making, they suggest, involves less prefrontal cortical activity. As –real life" involves adapting responses to in vivo circumstances, the prefrontal cortex is critical for this type of decision making. Thus, inhibiting in real life may require significant executive functioning capacities, whereas inhibiting responses on the NEPSY may not.

Literature Review

Conceptual Models of Executive Functions

As noted previously, when taken en masse, the models of executive functioning have as their common core the self-regulatory capacities that direct cognitive, emotional, and motor behavior in multiple settings. Since Luria‘s conceptualization, there has been increased discussion relative to the specificity, scope, and impact of executive functions. Denckla (1996) offered a definition of executive functions as a –domain-general" (p. 263) control and regulatory capacity, but cautioned against extending this concept to control of emotions and motivation as well as to assigning executive functions a higher-order power—similar to the intelligence theorists' construct of *g*. She later noted that defining the concept of executive functions poses a difficulty because of the ~~‗b~~readth of functions and developmental dynamics" (2007, p. 9) that can legitimately be included under the executive function umbrella, suggesting an evolution in thinking which increases the specificity of the functions. Recent delineations provided by others in the fields of psychology, neuropsychology, and education have worked to provide more conceptual specificity regarding what should be encompassed under this global term. As noted previously, Gioia et al. (1996), in their development of the Behavioral Rating Inventory of Executive Functions (BRIEF), describe eight executive subdomains: inhibition, shifting, emotional control, (thought to be part of a general domain labeled behavioral regulation) and initiation, planning/organization, organization of materials, working memory and self-monitoring (thought to be part of a general domain labeled metacognition). McCloskey et al. (2009) expand the knowledge concerning the specific subdomain executive functions, enumerating twenty-three distinct self-regulation executive functions and discussing how these functions interact with, through signaling or directing the cueing of, the cognitive abilities of reasoning, language, visuospatial and memory capacities. McCloskey posits the idea that these 23 executive function capacities differentially regulate functioning within four general domains; these are perception, cognition, emotion and action, which are employed differentially across four different arenas of involvement: interpersonal regulation, intrapersonal regulation, regulation of self within the physical world, and regulation of activities involving symbols or systems (such as reading, writing, and mathematics).

Although there has been greater elaboration in the models of executive functions and the impact and scope of these capacities, the issue of whether or not executive functions are truly dissociable from cognitive abilities, particularly fluid reasoning, remains a point of debate (Decker, Hill, & Dean, 2007; Denckla, 1996; Salthouse, 2005). Decker et al. note that by definition, executive functions and fluid reasoning are the same constructs, because both require the use of reasoning capacities to respond to novel situations or tasks. Denckla notes the complexities involved in attempting to separate executive functions from the fluid nature of –*g*,” but conversely provides data relative to the minimal relationship between executive functions and cognitive abilities (Reader, Harris, Schuerholz, & Denckla, 1994). Salthouse‘s investigation of the relationship between performance on executive function tests and on cognitive tests found a strong correlation in typical adults between executive function measures with reasoning and perceptual speed abilities. Executive functions, he concluded, may not be separable from the more basic cognitive abilities. Other research suggests that specific task demands determine the degree to which cognitive and executive capacities are dissociable; for example, functional neuroimaging of presumably typical adults during deductive reasoning tasks revealed significant activation in the left frontal and temporal lobes (with particularly strong frontal involvement) for reasoning in familiar situations, but a bilateral frontal-parietal, visual-spatial network activation when reasoning with unfamiliar material or in unfamiliar situations (Goel and Dolan, 2004). Still others (Konishi et al., 2002) note primarily frontal involvement in execution function tasks involving reasoning.

Executive Functions and Psychopathology

Issues of self-regulation and behavioral and/or emotional control appear particularly problematic in most manifestations of psychopathology; there is, indeed, an abundance of literature which discusses self-regulation and executive function deficits in many mental disorders. Studies of ADHD, OCD, autistic and depressed probands generally have identified the existence of executive control deficits (Barkley, 1997; Lawrence et al., 2006; Pennington and Ozonoff, 1996). By virtue of their definitions and criteria for diagnosis, behavioral disorders, particularly those known as the "disruptive behavior disorders," involve difficulties with regulation of behavior and self-control. Diagnostic criteria (*Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision;* American Psychiatric Association, 2000) for Attention Deficit/Hyperactivity Disorder (ADHD) include difficulties with sustaining attention and mental effort, with task completion, inhibiting response, following rules, and with regulating verbalization and motor activity. Oppositional Defiant Disorder (ODD) involves difficulties with control of temper, verbal regulation, and other aspects of emotional regulation (DSM-IV-TR). Diagnostic criteria for Conduct Disorder (CD) also include difficulties with following rules and demonstration of behavioral control according to societal norms, although implied in the diagnosis of CD is the deliberate nature of the behaviors. Other developmental or psychiatric disorders also involve self-regulation difficulties. Autism and related disorders are manifested by "restricted repetitive and stereotyped patterns of behavior" (DSM-IV-TR, p. 75). Those with depression often fail to sustain energy and effort for task completion or fail to initiate an activity or task altogether. Tourette's Disorder results in difficulties with motor and/or verbal inhibition. Given the fact that a diagnosis for these disorders is not determined unless the behaviors are significantly different from that which would be expected, based on the individual's mental age (cognitive functioning), it would appear that these difficulties with behavioral regulation and self-control might stem, at least in part, from deficits in frontal lobe functioning; that is, that deficits in executive functions are in large part the contributing factors to the outward manifestations of dysregulated behaviors. However, there is considerable discussion in the literature about whether or not these deficits are primary, causal factors in the disorder, an artifact of the clinical assessment, or a secondary effect of the disorder (Pennington & Ozonoff). Even less clear, given the emerging belief that executive functions are not a unitary function, but a collection of dissociable control processes, is the understanding of those specific neuroanatomical structures and executive function deficits that are involved in the dysregulation of behavior as evidenced in these disorders.

Relationship of Brain Functioning and Executive Functions

As noted above, although the relationship of brain functioning and behavior has been long discussed, neuroimaging advances over the past few decades have brought increasing specificity to the discussion of brain/behavior relationships, especially in the case of executive function capacities. As a significant refinement of Gioia et al.'s ―frontal system," Lichter and Cummings (2001) describe five frontal-subcortical circuits, three of which appear critically important to a broader conceptualization of executive functions. In addition to motor circuit and oculomotor frontal lobe circuits, Lichter and

Cummings describe a dorsolateral prefrontal circuit (DLPF) which mediates planning, organizing, use of feedback, working memory and retrieval executive functions (those which have been referred to traditionally as core ―executive functions”); the orbitofrontal (OF) circuit, which mediates behavioral inhibition and judgment capacities; and the anterior cingulate (AC) circuit, which mediates motivation, monitoring, and initiative. They suggest that each circuit is just that, a loop which interconnects with other regions of the brain—a closed loop which receives and transmits dedicated neurons and an open loop which is able to receive noncircuit information. Each circuit, by different pathways, routes through the striatum and globus pallidus, to the thalamus and then back to the frontal cortex. Other literature (Middleton & Strick, 2000) suggests similar ―loop” circuitry from the basal ganglia to the prefrontal cortex, disturbances which appear to underlie specific neurological and psychiatric (i.e. cognitive and behavioral) symptoms. It may be this connectivity and interconnectivity which contributes to the complexity of the discussion relative to whether or not the frontal/executive function components of each circuit are causal or merely contributory. Although the structure that Lichter and Cummings describe would suggest distinct, separate circuitry for cognitive and behavioral inhibition, the body of literature as a whole suggests that the complex interactions of neural pathways do not appear to result completely or easily in dissociable mechanisms for what is observed in the cognitive and behavioral realms.

Neuroimaging advancements appear to support the structural and/or functional differences in the frontal and subcortical areas of the brain in most psychiatric disorders. As noted by Pennington and Ozonoff (1996), ―the best evidence for differences in ADHD comes from measures of brain function;" the authors posit the idea that reduced right frontal blood flow is an indicator of decreased executive controls. The literature points to anomalies in the anterior cingulate, dorsolateral and orbitofrontal circuits (Starkstein & Kramer, 2001; Voeller, 2001) and right prefrontal striatal circuitry (Barkley, 1997) in subtypes of ADHD. Literature relative to autism suggests involvement of the DLPF, OF and AC circuitries (Klin, Volkmar, & Sparrow, 2000) and reduced right frontal blood flow (Pennington & Ozonoff); however, there continues to be much discussion regarding the physiology of autism (Lichter & Cummings, 2001). Mayberg (2001) reports on literature on depression, noting decreased frontal lobe function, specifically in the dorsal, ventral and rostral region, implicating the DLPF and AC circuits. Brain imaging in patients with Obsessive-Compulsive Disorder (OCD) is suggestive of abnormally high activity in the orbital prefrontal cortex (Baxter, Clark, Iqbal, & Ackermann, 2001). Hale, Blaine-Halpern, and Beakley (2007) hypothesize that the over- and under-activity of cortical-subcortical circuitry mediates the manifestation of the executive function difficulties. Underactivation of these circuits leads to distractibility, impulsivity, and hyperactivity, whereas overactivation leads to fixated and repetitive behaviors and hypoactivity. Although there is not conclusiveness relative to whether or not these functional or structural differences are the causal factors in psychopathology, there is certainly widespread agreement about the existence of frontal-subcortical involvement in these disorders.

Neuropsychological Assessment of Executive Functions

Neuropsychological assessment data, particularly from measures of executive functions, have long been used to understand the impact of deficits resulting from frontal lobe damage or injury. Although the previously cited literature does not indicate specific brain injury in those with different psychopathology, it does suggest frontal-subcortical anomalies. It is then not surprising that the results of neuropsychological assessments suggest specific impaired executive function capacities in those with psychopathology. This is especially true in ADHD and ADHD subtypes. In his review of neuropsychological assessment findings for ADHD children, Barkley (1997) notes the following deficits: inhibition, emotional control, verbal working memory, time awareness and planning, verbal fluency and response flexibility, and motor control and sequencing. Other findings support deficits in inhibition in ADHD-Hyperactive/Impulsive subtype (Pennington, 2002); initiating, sustaining, planning, organization, and working memory in the ADHD—Inattentive subtype; and inhibiting, shifting, self-monitoring, and emotional control in the combined subtype (Gioia et al., 2001). A meta-analytic review (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005), however, was less clear on the executive function deficits by ADHD sub-type. Although the review found significant executive function impairment in ADHD groups specifically related to response inhibition, vigilance, working memory, and planning, these deficits were evident in the Inattentive and Combined Type subgroups, but not Hyperactive/Impulsive subtype. Although the authors noted the small size of this subgroup, they did suggest that ADHD-HI may have an etiology different from the other two subtypes.

Not surprisingly, given its diagnostic criteria, autistic spectrum research has identified deficits in cognitive flexibility, planning, verbal fluency, inhibition, and interference control (Verte, Geurts, Roeyers, Oosterlaan, & Sargeant, 2006); working memory (Goldberg et al., 2005); and cognitive flexibility and planning (Lopez, Lincoln, Ozonoff, & Lai, 2005; Ozonoff et al., 2004). Although there appears to be a paucity of literature relative to assessment of the executive functions of depressed individuals, that which exists is suggestive of deficits in self-correcting (Holmes & Pizzagalli, 2007) and verbal fluency, inhibition, set-maintenance, and working memory (Stordal, 2006); in those with Bipolar II Disorder, deficits include impaired working memory and set-shifting (Torrent et al., 2006). Decision-making and set-shifting impairments were found in OCD probands (Lawrence et al., 2006).

As mentioned, these previous findings are likely to be based on clinical interviews as well as performance on standardized measures and/or behavior ratings which are developed to assess executive function capacities. Neuropsychological assessment is thought to be another way of examining brain functioning because behavior is presumed to be the outward or overt manifestation of brain functioning (Lezak, Howieson, & Loring, 2004). Traditional neuropsychological test batteries, such as the Halstead-Reitan Neurological Test Battery, the Luria-Nebraska NTB, and Stroop Color-Word Test, the Tower tests—London, Hanoi, and Toronto, and the Wisconsin Card Sorting Test have been in existence for decades. These and more recent measures such as the D-KEFS and the BRIEF, mentioned previously, have found their way into school and clinical settings because of their development and/or modifications for use in assessing children. Most of the tasks designed to measure executive functions on these newer batteries are very similar to the tasks found on the older, traditional neuropsychological instruments.

The Wisconsin Card Sorting Test

WCST Design and Constructs.

The Wisconsin Card Sorting Test has been deemed the most widely used of all the neuropsychological assessment tools for evaluating executive function capacities (Kaplan et al., 2006; Konishi et al., 1999; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Originally developed by Grant and Berg as a test of abstract reasoning for typical adults, it is now used primarily as neuropsychological instrument to assess the integrity of frontal lobe functions (Heaton et al., 1993). The 1993 Heaton version was standardized and normed for use with children and adolescents, beginning at the age of 6 ½ years, with a total sample size was 899. Thus, the WCST can be used with a wide age range of individuals, from 6 ½ to 89, although the authors provide a cautionary note about performance interpretation for the 85-89 age group because of the very small size of individuals in the normative group in that age category.

The WCST is different from other ―tests” of executive functioning capacities because it provides relatively ambiguous directions for completion which puts the onus for executive self-cueing on the examinee. As one can observe from the directions given in the manual, the examinee is given little information relative to that which constitutes successful completion of this task:

> This test is a little unusual because I am not allowed to tell you very much about how to do it. You will be asked to match each of the cards in these decks to one of these four key cards. You must always take the top card from the deck and place it below the key card you think it matches. I cannot tell you how to match the cards, but I will tell you each time whether you are right or wrong… (Heaton et al., 1993, p. 5).

Following these directions, the examinee has to match 128 stimulus cards to the four key cards. The first matching principle is that of color. After the examinee has made ten consecutive color matches, the matching pattern is switched, without warning, to the matching principle of the shape of the symbols on the cards. After ten consecutive, correct symbol matches are made by the examinee, the pattern is switched again – this time to the number of symbols on the card. Ten successive, correct matches are again required, after which the matching principle is switched back to color for a second series of set switching by color, symbol type and number. Thus, although the WCST is an assessment of problem-solving, shifting and inhibiting (because the matching principle changes without prior warning), sustaining of attention and effort, and use of working memory, the examinee is not cued about the need to engage these executive capacities; he or she must independently recognize the need to use them.

WCST performance is quantified along several dimensions and scores are represented as standard scores (mean = 100; standard deviation of 15) and/or percentiles, based on age and, for adults, age and educational level. Scoring areas include number and percent of errors, number and percent of perseverative responses, number and percent of perseveration errors, number and percent of nonperseverative errors, and percent of conceptual level responses (conceptual responses are correct responses that occurred in sequences of three or more). Total errors represent the number of times that the examinee is told that his or her card selection is ―wrong,” even when the test is shifting to a new matching dimension. Perseverative responses represent the number of times that the examinee maintains a previous matching pattern even when told that that pattern is incorrect. Perseverative responses can be unambiguous—representing a clearly ―stuck-in-set” response—or ambiguous—when the card matches for a new (and possibly correct) dimension, but also matches for the previous, now incorrect, dimension. Nonperseverative errors are all those remaining errors that do not fit the perseverative description. Conceptual level responses, as noted, are those sequences of three or more correct responses that suggest the examinee has ascertained the matching pattern. Standard and percentile scores are provided for these scoring areas. Percentile ranges only are given for other scoring areas: number of categories completed, trials to complete first category, failure to maintain set (defined as an incorrect response after providing five or more successive correct responses), and a learning to learn dimension. This last measure is designed to assess the examinee‘s average change in performance (i.e., improved efficiency) as he or she progresses through the completion of the categories.

As noted previously, although initially constructed to assess abstract reasoning ability and cognitive flexibility (defined as shifting cognitive strategies), the WCST is now used to measure strategic planning and organizing, directing behavior toward a goal, utilizing feedback and set shifting, inhibiting impulsive responding, and the employment of working memory capacities—many of the self-regulation capacities that are central to the more elaborative models of executive functions that have been more recently proposed. In their review of the literature on executive functions, Alvarez and Emory (2006) cite the following as the underlying executive components of the WCST: inhibition and switching, working memory, and sustained and selective attention. In a factor analytic study of the WCST, Greve, Stickle, Love, Bianchini, and Stanford (2005) cite three processes: set-shifting, problem solving or hypothesis testing, and response maintenance. Miyake et al. (2000) identified three dissociable factors of the WCST—inhibition, shifting, and updating (a working memory function), although more recently these same researchers (Godinez et al., 2005) suggested that updating was not significantly related to WCST score indices. Because the literature tends to cite all these components as dissociable dimensions of executive function control, use of the WCST as an assessment of executive function capacities is well-documented (Barcelo, 2001; Greve et al.; Watson, 2005).

Much of the abundant literature on the WCST has focused on the processes that underlie performance as well as on the neuroanatomical structures, both frontal and nonfrontal, that are activated in the test performance of adults. Consistent with findings suggesting that the WCST assesses executive functions are neuroimaging studies that pinpoint specific regions of the prefrontal cortex, specifically the orbito-frontal, frontal-striatal, dorsolateral prefrontal, and inferior prefrontal regions of the cortex, that are activated during WCST performance (Lie, Specht, Marshall, & Fink, 2005; Nagahama et al., 2001; Konishi et al., 1999; Seidman, Faraone, Biederman, Weber, & Ouellette, 1997; Kimberg and Farah, 1993). Lie et al. describe a WCST functional network which involves activation of right dorsolateral prefrontal and anterior cingulate regions. A primary involvement of the dorsolateral prefrontal cortex is also noted by Alvarez and Emory, but they too cite the apparent activation in the orbitofrontal and ventromedial cortices. This particular finding has special relevance for this study because the dorsolateral region is believed to mediate more metacognition (internal) activities but the OC and AC appear to be involved in more behavioral (observable) regulation, suggesting a confluence of cognitive and behavioral control processes for successful task completion. Alvarez and Emory (2006) note:

> A parsimonious explanation of the WCST results supports the idea that a distributed network of neural circuits is activated when task demands involve integrated functioning. For example, activities of daily living, such as planning a trip to the store, involve overt and covert behavior components. At the overt level, the individual may search for the appropriate writing instruments, write down directions, and make a list of items. At the covert level, the individual may engage long-term and short-term memory functions, visualize a path to the store and where items are located, and plan a budget that is within the parameters of the resources available. One could refer to these activities as internal and external (or implicit and explicit) representations of cognitive abilities that fall within the purview of executive functions (p. 22).

This complexity is also evident in the studies of WCST performance in terms of the brain-basis for the specific executive function capacities thought to be needed for successful task completion. There is considerable discussion regarding the lack of clarity or specificity of the constructs or capacities required for task success or for the way in which these capacities are represented in the WCST scores (Alvarez & Emory, 2006; Barcelo, 2001; Greve et al., 2005; Miyake et al., 2000). This is especially true in the studies of set-shifting capacities/perseveration, even though there appears to be general agreement that set-shifting is the hallmark of the WCST (Barcelo; Miyake et al.).

As noted previously, the ability to inhibit perseverative responding and shift set are considered critical to successful performance. A review of the literature suggests different perceptions regarding how perseverative behavior and shifting capacities may be defined. As noted previously, there appears to be more than one dimension of perseveration that is tapped with the WCST: the stuck-in-set perseveration in which the examinee continues to respond to the prior matching pattern in spite of being provided with feedback about the inaccuracy of responses; the perseverative responding in which the individual may be shifting to a new matching principle, yet it is not evident because of the confounding qualities of the card (i.e., that it may match for more than one principle at any given time); and perseverative responding in which there is a return to the previous matching principle when other matching attempts have not been successful.

These different types of perseveration may in fact activate different regions of the brain (Nagahama, Okina, Suzuki, Nabatame, & Matsuda, 2005; Stuss et al., 2000). Neuroimaging of the first type of perseverative responding (stuck-in-set) suggests activation in the bilateral rostrodorsal PFC and the left frontopolar cortex, confirming other studies suggesting the same activated areas. In contrast, in the case of perseverative responding of the third type (return to a prior pattern), there was not frontral, but posterior lobe activation, specifically in the left parietal lobe (Nagahama et al.). In reference to what may be the more surprising findings of this study, i.e. the implication of the parietal regions in the return to a prior set, the authors note that one of the roles of the parietal lobe is visual attention, a likely underlying component necessary for achieving success with the WCST. Konishi et al. (2002) dismantle the set-shifting capacity in a somewhat different way, seeing it as two components, one of exposure to negative feedback (i.e., being told that the matching attempt was incorrect, because of the unannounced shift in pattern) and the other of updating to a new set or matching dimension. With these definitions of set-shifting, they found activation in three bilateral frontal regions during set-shifting on the WCST; the right frontal hemisphere was activated when the examinee experienced negative feedback to his or her matching attempts and the left frontal region was activated when the correcting mechanism (changing to a new set) kicked in. An even more recent study, however, maintains that the right frontal lobe is critical for set-shifting capacities, combining the capacity of responding to the negative feedback with the new shifting response, yet not ruling out involvement of either the left frontal or parietal lobe (Carillo-de-la-Pena & Garcia-Larrea, 2007). Thus, it is quite instructive to note the extent to which the literature highlights the degree to which the findings mirror the researchers' viewpoints about how the executive function capacities are defined.

Although less research has been done exploring WCST performance and working memory, a similar diversity of findings is observed. On the surface, working memory capacities appear to be required in WCST performance; it seems reasonable to posit the idea that the examinee needs to hold in mind knowledge of the prevailing, as well as the prior, patterns required for correct responding. Some suggest that the same inferior prefrontal area is implicated with both set-shifting and working memory in performance on the WCST (Konishi et al., 1999); other literature suggests a more diffuse interaction of frontal circuits for maintaining associations in working memory (Kimberg & Farrah, 1993). Still others suggest that there is less evidence of a working memory component when performance on the test is compared with other measures of working memory (Stratta et al., 1997). The diversity of opinions here, as was observed in the literature on the capacity to shift set, may center on the identification of the specific capacities thought to be involved in the engagement of working memory because some note that working memory may involve multiple components, such as central attention control and verbal/nonverbal feedback loops (Baddeley, Chincotta, & Adlam, 2001).

Differentiating working memory from attentional capacities also appears problematic in attempts to understand those cognitive capacities that contribute to WCST performance. In the literature addressing attention and WCST performance, there is considerable discussion regarding the definition and measurement of attention. Most researchers seem to agree that Failure to Maintain Set is an indication of loss of sustained attention; however, assessment of this must be disassociated from the task requirement of initially reaching set maintenance: One cannot be found to be lacking in attention until attention has in fact been directed to the goal (Greve, Williams, Haas, Littell, & Reinoso, 1996). Thus, measurement of inattention, according to Greve et al., can be conducted only when the problem-solving factor is eliminated. Additionally, there is some disagreement about the adequacy of the Heaton definition and scoring of Failure to Maintain Set (in which set loss is indicated only after 5 consecutive correct responses have been provided) because it could be argued that set loss may legitimately be established with fewer consecutive correct responses preceding the incorrect response (Stuss et al., 2000). For example, when defining loss of set as an error occurring after only three consecutive correct responses, findings suggested that activation of the inferior medial prefrontal cortex and orbitofrontal regions contributed to successful maintenance of set (Stuss et al.)

WCST and Populations Studied.

The relationship between performance on the Wisconsin Card Sorting Test and traumatic brain injury has been studied almost since the inception of the test. Poor performance on the WCST has been linked to damage to the dorsolateral prefrontal cortex (Kimberg & Farah, 1993). More specifically, Kimberg and Farah indicate that the tendency to perseverate (that is, not shift set when the task dictates) is the primary deficit underlying poor performance on the WCST, although they go on to note that working memory impairments resulting from frontal lobe damage may also contribute to poor performance. However, in spite of what looks like fairly compelling evidence for the WCST as a measure of frontal lobe damage, Alvarez and Emory‘s meta-analytic review noted that in some studies WCST performance failed to differentiate between frontal and non-frontal damage and suggest that the test has sensitivity to, but not specificity for, frontal lobe damage. Heaton et al. (2000) in the WCST manual also noted that individuals with frontal lesions are only slightly more likely to demonstrate poorer WCST performance than those with more diffuse cerebral damage.

Less clear and at times more contradictory are studies that relate performance on the WCST with less distinct forms of frontal damage, but rather, with those that demonstrate the behavioral manifestations of weak executive control. In his review of the existing literature relative to WCST performance, Barkley (1997) noted that ADHD children demonstrated difficulties relative to perseverative responding despite the verbal feedback relative to their incorrect responses. Other studies (e.g., Willcutt et al., 2005) found no difference in perseverative errors between ADHD and control subjects. In their study of 118 male subjects diagnosed with ADHD and 99 male controls, Seidman et al. (1997) found that the ADHD probands were significantly more impaired on the perseverative and non-perseverative error scores on the WCST, suggesting the possibility of deficits in shifting, inhibiting, and/or working memory. It is important to note, however, that many of the ADHD probands in this study were taking psychostimulant and possibly other medications, which may or may not have impacted WCST performance. That medication response might affect performance is suggested by a study by Hale, Fiorello, and Brown (2005) which found that the children diagnosed with ADHD, who demonstrated significant deficits in executive functions (as measured in part by WCST performance) were most likely to show behavioral and cognitive improvements when taking medication. In their study of individuals with frontal damage who demonstrated “behavioral abnormalities in their everyday behavior” (1997, p. 189), Baddeley, Della Sala, Papagno, and Spinnler (1997) found no difference in the WCST performance of this group versus a non-dysexecutive group, although the researchers did employ a shorter-version of the test, requiring only that the first three of the six categories be completed. This alteration of format, although standardized, may represent a significant change in task complexity, because the responsible investigator has noted that many individuals are able to respond to the initial shifts from color to form to number, but encounter significant difficulties when required to make the shift back to the initial concept of color.

The autism literature also reports on significant difficulties in performance on the WCST, particularly as it relates to perseverative responding. Lopez et al. (2005), who evaluated 17 adults with autism, found that they completed significantly fewer matching patterns, made more errors, and generated more perseverative responses than controls. Verte et al. (2006) also found more perseverative errors on the WCST among high-functioning autistic and PDD-NOS (Pervasive Developmental Disorder-Not Otherwise Specified) probands. It is interesting to note, however, that the autism literature also suggests that these deficits do not always carry over to the computerized version of the WCST, with the implications being that the computerized version might reduce some of the verbal and social aspects of task engagement (Ozonoff, 1995), and as a result, may activate different brain structures.

Not surprisingly, given the patterns seen above, poor WCST performance is also found in other psychopathology, including depressive disorders (Torrent et al., 2006; Borkowska & Rybakowski, 2001), OCD (Lawrence et al., 2006), and schizophrenia (Demakis, 2003; Heaton et al., 1993).

In this body of literature regarding WCST performance and the populations that experience difficulty, there are very few references to results relating to child and adolescent performance, even though the test was standardized and normed for children fifteen years ago. Prior to this development, Chelune and Thompson (1987) discussed the value of neuropsychological assessment of children, not to identify the loss of neuropsychological function, as the test had been utilized to that point for adults, but to understand skill acquisition, i.e. to understand the emergence and development of those executive function capacities in children. An electronic search of the literature (PsychInfo/Ovid, Medline) yielded fewer than twenty-five journal articles or dissertations that focused specifically on WCST results of children and adolescents.

Heaton et al. (1993), in their review of the variables from the standardization sample, suggest that three constructs measured by the WCST may discriminate clinical groups from each other and from normal children: the ability to think abstractly and shift cognitive set (as measured by Error Percentage, Perseverative Responses, Conceptual Level Response Percentage and Trials to Complete First Category); the ability to maintain cognitive set (Error Percentage and Conceptual Level Responses), and the ability to demonstrate flexibility in thinking (Perseverative Responses, Categories Completed, and Trials to Complete First Category). A meta-analysis of the test's sensitivity and specificity (Romine et al., 2003) found medium effect sizes in ADHD children for four WCST variables: Percent Correct, Total Errors, Number of Categories, and Perseverative Errors. In children with learning disabilities, large effect sizes were found on three variables: Number of Categories, Total Errors, and Non-perseverative Errors. Although the results of studies varied, relative to control group performance, adolescents with conduct disorder showed medium effect sizes for Perseverative Errors and Categories Completed and large effect sizes for Total Errors. Large effect sizes in Perseverative Responses and Errors were generally found in children with an autism spectrum disorder. Children with anxiety disorders were also found to have more

Perseverative Responses and Total Errors than controls; as was the case with studies of adults with depression and executive/WCST difficulties, there was very little literature available on children and adolescents with depression. In conclusion, the authors note that a poor performance on the WCST is likely indicative of a disorder; however, the results do not contribute either to a differential diagnosis nor to the fact that they are likely to be helpful, in younger children especially, in understanding the specific underlying executive function deficits (Romine et al.). Barkley, Grodinsky, and DuPaul (1992), in their review of 22 studies of children and adolescents with ADHD (with and without hyperactivity) found no differences in ADD subtypes or in comparisons with non-ADHD individuals overall, but they did note that WCST performance improved with age and suggested that the extent of impairment may be lessened as children go through their teen years.

A review of the literature suggests that WCST scores themselves may not be adequate for describing a child‘s performance on the WCST and that application of a process approach in WCST interpretation may yield important information relative to an individual‘s cognitive and executive functioning capacities. The process approach to administration and assessment distinguishes process from product (scores), posits that test scores may not reflect the same underlying single construct in all individuals, and considers multiple variables, both quantitative and qualitative, that contribute to test performance (Kaplan, 1990; McCloskey & Maerlander, 2005). McCloskey and Maerlander note that complex tasks such as those found on cognitive (and presumably) executive function assessment instruments involve the interactions of multiple neuropsychological processes and that, in fact, any one or more of these individual processes may be at the root of poor test performance. Relative to the WCST specifically, Heaton et al. (1993) indicated that WCST scores themselves may not offer sufficient information for a clinical diagnosis in children and further suggested that the ―relative pattern of performance" (p. 54) may be helpful in identifying executive function difficulties in children and adolescents. Denckla (1996) goes further in stating ―without error analyses as the bases for generating a distinctive pattern of failure (more than just a low level of quantitative score), the WCST is declared nonspecific for either EF (sic) or frontal injury" (p. 269).

As has been previously noted, there is an abundance of studies involving the WCST, the executive capacities that it purportedly measures, and the relationship between WCST performance and psychopathology. In spite of this, a definitive and specific understanding of what is involved in performance on this test remains elusive. Numerous efforts to create better understanding by modifying or shortening the original and Heaton versions of the test have, in all likelihood, served to enhance specificity for that particular modification, but resulted in weaker generalizability to the test itself. Although most studies, particularly those which involve neuroimaging, involve use of the computerized version of the WCST, there is little data (except with the aforementioned autism research) that explores the similarities or differences in test results based on administration format. Thus, because of the complexities and ambiguities of the test, definitive causal factors appear to have been difficult to delineate and further study is necessary As noted by Stuss et al. (2000), ―while such a complex multifactorial test as the WCST is unlikely to be sensitive only to the functions of the frontal lobe, analysis of the cognitive processes involved can be helpful in understanding why and how individuals with lesions in different brain regions may be impaired on this test" (p. 388).

Delis-Kaplan Executive Function System

Although it is most frequently used, the WCST is not the only test that is used to assess executive function capacities. Other instruments purport to measure constructs similar to those on the Wisconsin Card Sorting Test. For example, the Delis-Kaplan Executive Function System (D-KEFS; Delis et al., 2001), normed for individuals 8-89, provides nine different executive function tests, most of which mirror those found in older neuropsychological assessments. Each test or task stands alone, with no ―Executive Function" composite score provided. Tasks involve executive function involvement with orally presented verbal information and visually presented nonverbal and verbal information. . The nine tests are: Trail-Making Test, Verbal Fluency (letter and category), Design Fluency, Color-Word Interference Test (similar to the Stroop test), Sorting Test (as described by the authors, a game-like Wisconsin Card Sorting Test), Twenty Questions Test, Tower Test (a modified version of previous Tower tests), and Proverb Test (for individuals 16 and older). Standard scores, with a mean of 10 and a standard deviation of 3, are derived for each test and there are multiple primary measures for fluency, shifting, inhibition, etc., with scoring provided for optional measures. These optional measures, specifically those called contrast scores, often compare the performance on more fundamental or skill/ability-based tasks such as verbal fluency with the performance that requires shifting or inhibiting capacities within the task. Although some question the utility of so many scoring options (Lezak et al., 2004), the authors of the D-KEFS discuss the importance of the cognitive-process approach as a way of capturing the multiple dimensions and complexities of executive function capacities (Delis et al.). As they note, –one of the challenges in the development of executive-function tests is to provide, as much as possible, rigorous empirical means for determining whether poor performance is due to deficits in more fundamental cognitive skills or deficits in higher-level executive functions” (p. 3).

Because of its relatively recent publication date, there is minimal data regarding the brain-basis for D-KEFS performance or the populations that demonstrate difficulties with these tasks. Also, the authors provide no validation data for the specific subtests of the D-KEFS. However, given the tests‘ similarities to other, older tests of executive functions, there is considerable support for some of the D-KEFS tasks as measures of executive function capacities and evidence that the brain-basis for performance rests at least in part in the frontal regions of the brain. For example, performance on the Trail-Making Number-Letter Switching (requiring shifting of set between numbers and letters to draw a trail sequence), like its Trail-Making Test Part B predecessor, appears sensitive to dorsolateral frontal damage (Stuss, Floden, Alexander, Levine, & Katz, 2001). Verbal fluency tests, because of their retrieval, self-monitoring, self-initiation, and inhibition requirements, are often used to assess executive function difficulties and have been found to be indicative of frontal lobe activation (Henry & Crawford, 2004); fMRI studies suggest activation in superior frontal sulcus for word generation (Phelps, Hyder, Blamire, & Shulman, 1997) and the left inferior frontal gyrus for word switching (Hirshorn &

Thompson-Schill, 2005). Design fluency tasks, although they measure use of visualmotor capacities, are considered to be similar to the verbal fluency tasks, use of which has been well-established (Baron, 2004; Salthouse, 2005). Right and left frontal activation appear to be necessary for D-KEFS Design Fluency tasks (Baldo, Shimamura, Delis, Kramer, & Kaplan, 2001). There is less literature on tasks such as the Twenty Questions Test and Word Context Test, although forms of these tasks have existed for some time. As noted previously, the Tower Test is similar to previous, well-documented versions, although the validity of its use with children is in question (Baron). Performance on the Stroop Color-Word Test is impacted by the left dorsolateral frontal lobe, bilateral superior medial frontal, and possibly, the anterior cingulate cortex (Stuss et al., 2001), so it is logical to assume that the same neuroanatomical structures are activated with performance on the D-KEFS Color-Word Interference Test.

Although many of the tasks of the D-KEFS would not be considered as comparable measures of executive functions with the WCST because of their strongly verbal involvement, the additional components of inhibition and switching constructs in these tasks invites comparisons between the two instruments. Additionally, the WCST does not appear to be intrinsically a verbal task; however, it is possible that some examinees may be accessing verbal capacities when engaged in task completion. Thus, a comparison of the two instruments may not provide information regarding the neural activation or specific executive/cognitive capacities that are engaged in completion, but it is still warranted. In fact, Delis, Kaplan, and Kramer provide correlational data on two assessment tools, the California Verbal Learning Test, Second Edition (CVLT; Delis, Kaplan, Kramer, & Ober, 2000), with which there generally were no or low positive correlations (Delis et al., 2001), and the WCST. Although the numbers used in this study were small (*N* = 23), there were moderate positive (range of .44 to .59) correlations between the D-KEFS Sorting Test and the WCST Categories Completed; additionally, WCST Categories Completed had a low to moderate correlation with many other D-KEFS achievement scores. Interestingly, WCST Perseverative Responses had a low positive correlation with D-KEFS Color-Word Interference Inhibition (.23) and Inhibition/Switching (.20) and Trail-Making Test Number-Letter Switching (.44); however, the correlations between WCST Perseverative Responses and other D-KEFS switching tasks (Design and Verbal Fluency) were negative (Delis et al.). Although the authors note some apparent degree of shared variance, stating that further study with larger samples need to be completed, they also indicate that each instrument offers unique variance in the assessment of executive functions.

Instruments such as the D-KEFS measure an individual‘s executive functions; however, they may actually assess performance under which much external executive control is subsumed by the examiner because he or she may provide, as Gioia et al. note, ―the structure, organization, guidance, and plan, as well as the cueing and monitoring necessary for optimal performance" (2001, p. 338). Thus, the executive function demands are lessened in these ―executive" tasks. For example, although many of the D-KEFS subtests assess the individual‘s capacity to shift cognitive set (by alternating between semantic category in the verbal fluency task, switching alphabetic and numerical sequence in Trail-Making, etc.), the switch is not only cued, but practiced before the task begins. Additionally, the D-KEFS, in a manner similar to the WCST, may be measuring executive function capacities in only one arena—that of executive function capacities for tasks that involve manipulation of symbols.

Wechsler Intelligence Scale for Children—Fourth Edition and Wechsler Adult Intelligence Scale—Third Edition

Although the WCST was originated to assess abstract reasoning, this cognitive capacity is generally measured through other standardized assessments, such as the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV; Wechsler et al., 2004) or the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III; Wechsler, 1997). Both were designed to measure the cognitive components of intelligence through assessments of verbal comprehension and knowledge, perceptual organization, abstract reasoning, quantitative reasoning, memory, and processing speed. The Wechsler provides an overall global IQ score (Mean = 100; standard deviation +/- 15) and scores to measure indices of verbal comprehension, perceptual reasoning (or organization for the WAIS-III), working memory, and processing speed; scores (mean = 10; standard deviation = +/- 3) are also obtained for the more discrete cognitive functions that make up these indices (Wechsler).

In discussions of the WCST relative to the measurement of reasoning capacities, constructs such as fluid reasoning and abstract reasoning have been delineated as factors of the test. In the discussion of theoretical factors of the WISC provided in the test manuals (WISC-IV; Wechsler et al., 2004; WAIS-III; Wechsler, 1997), the subtests of Matrix Reasoning, Picture Completion, and Word Reasoning are considered measures of fluid reasoning, although the manuals note that working memory tasks, such as Letter-Number Sequencing, Arithmetic, and Digit Span Backwards, may also be indicators of fluid reasoning. Block Design requires fluid reasoning capacities; however, it also measures perceptual organization and, at younger ages, matching abilities (Wechsler, 2003; Weiss, Saklofske, & Prifitera, 2005), components also required in WCST performance. Because of its fluid reasoning demands and visual pattern stimuli, the components of Matrix Reasoning, especially, may offer an interesting comparison with elements of the WCST. For each item of the Matrix Reasoning Subtest, the individual is presented with an incomplete matrix and is asked to identify the missing component from 5 response options. Items are preceded with the verbal query: ―Which one here (with the examiner pointing to the 5 options) goes here (examiner points to the box with the question mark)?" (WISC-IV: Wechsler et al.). Not only are the verbal and nonverbal directions quite explicit, but the individual is also allowed 3 practice items before the actual test begins.

Although not designed to be tests of executive function capacities, other subtests of the Wechsler scales may assess those capacities. In their review of the Digit Span subtest, Hale, Hoeppner, and Fiorello (2002) found that Digit Span Backwards assessed the executive capacities of sustained attention, inhibition, sequencing, cognitive flexibility, and set maintenance and shifting. In addition to working memory tasks, those tasks that require processing speed (Coding, Symbol Search, and Cancellation) require the capacity to generate, gauge, pace, and monitor performance; Cancellation also requires the inhibition of response (Weiss et al., 2005). These are self-regulation executive function capacities in more current models of executive functions (e. g. McCloskey et al., 2009). In the WISC-IV Integrated (Wechsler et al., 2004), Elithorn mazes assess planning, self-monitoring, and the ability to inhibit impulsive responding (Weiss et al.). As one looks at the scope of the tasks on these latest versions of the Wechsler tests, it would be relatively easy to argue that some executive function capacities are required for effective completion on most of these subtests.

Although it is beyond the scope of this study to review the literature related to the brain basis for performance on the Wechsler instruments, it would appear that the posterior activation indicated on studies of WCST performance is, in part at least, due to the abstract reasoning, memory, and/or visual processing requirements of task completion. Certainly the abundance of literature relative to the lack of frontal lobe specificity and activation of regions other than the frontal lobes (Alvarez & Emory, 2006; Demekis, 2003; Salthouse, 2005) would suggest involvement of cognitive capacities that are not solely executive in nature. In the discussion of the validity of the WAIS-III, and specifically, of the correlations between the WAIS-III and the WCST, most, but not all, of these components were assessed. Using the scores of 21 clinical participants, performance on the two instruments was compared. Measures on the WCST of total correct, categories completed, total errors, and perseverative errors were correlated with the WAIS-III factors of Verbal IQ, Performance IQ, Full Scale IQ, and the four Index scores of Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed. The correlations suggested that the strongest relationship (.48) was between the WAIS-III Working Memory Index with the WCST total number correct.

WCST total number correct score was moderately related to Full Scale IQ (.42) and Perceptual Organization (.42) (Wechsler, 2002). Using another clinical population, correlations of the WAIS-R Full Scale IQ with total numbers of categories completed on the WCST yielded a stronger correlation (.77), although the author provides a cautionary note that the problem-solving skills, or lack thereof, may be an artifact of the particular clinical population studied (Heinrichs, 1990).

There is limited literature on the correlation between the WISC-IV and the WCST. Validity studies in neither of the test manuals cite any correlational data. The existing literature is often conflicting and limited by sample size or by generalizability of findings because the studies often involve specific clinical populations. One study (Riccio et al., 1994) noted stronger correlations, depending on the age of the children; another suggests that the higher the cognitive functioning, the stronger the performance in all categories of the test in typical children, although giftedness helped performance only in the younger group (Arffa, Loveli, Podell, & Goldberg, 1998). Children with identified nonverbal learning disabilities (with decrement visual-spatial-organization and problem-solving abilities) fare less well than those with verbally-based disabilities on most categories of the WCST (Fisher, DeLuca, & Rourke, 1997). Denckla (1996) notes a possible correlation with Verbal IQ in children. There appears to be minimal data relative to the correlations between WISC Perceptual Reasoning or Working Memory index scores or the specific subtests within these indices with WCST scores. Additionally, in line with Salthouse‘s argument that WCST performance better correlates with reasoning and processing speed capacities, there is no data relative to the Processing Speed index with WCST scores.

Behavior Rating Scales

Behavior Rating Inventory of Executive Functions.

Indirect measures such as behavior rating scales that focus on executive function capacities have also increased over the past decade. The mostly frequently used appears to be the Behavior Rating Inventory of Executive Functions (BRIEF), developed in 1996, as two questionnaires, one for parents and one for teachers of children five to eighteen. Gioia, Espy, and Isquith (1996) developed a BRIEF rating scale for the preschool (ages 2-5.11); a self-report form for children aged 11-18 was also developed (Guy, Isquith, & Gioia, 1996). The original BRIEF (teacher and parent questionnaires for ages 5-18) contains 86 questions, each rated with a 3-point scale to reflect the frequency of occurrence of the specific behavior (1 = Never, 2 = Sometimes, 3 = Often). Although the questions are organized in random fashion, the responses given suggest the degree to which an individual demonstrates, in his or her behavior, any or all of the eight executive function deficits noted previously. Response weights are tallied and converted into t-scores and percentiles, with the authors noting clinical significance for t-scores that equal or exceed 65. For example, an individual who receives a high t-score on scores on the Inhibit Scale is likely to have received ratings of 2‘s or 3‘s for several of the items of the Inhibit scale, which includes items such as ―interrupt others,” ―acts too wild or out of control,” and ―has trouble putting the brakes on his/her actions.” Likewise, a clinically significant t-score on the BRIEF Parent Form is likely to be the product of item ratings of *sometimes* and/or *often* for several behaviors such as "resists or has trouble accepting a different way to solve a problem with schoolwork, friends, chores, etc.," "acts upset by a change in plans," and "tries the same approach to a problem over and over even when it does not work."

Behavior Assessment System for Children, Second Edition (BASC-2).

Although the Behavior Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004) now contains a supplemental content scale labeled Executive Functioning, the scale comprises only 5 items drawn from various other BASC-2 subscales. Overlooked in the development of the BASC-2 is the fact that many of the items on the BASC-2 assess behaviors reflecting either the ineffective, or effective, use of executive function capacities. This should not be surprising, given the fact that the hallmark of many diagnoses of emotional and behavioral difficulties involves problems with self-regulation. The BASC-2 is described as a multimethod tool to describe and evaluate behavior and perceptions of children and adolescents from the ages of 2 to 25. It consists of two rating scales, one for parents and one for teachers, and a self-report form for children, adolescents and young adults aged 8 to 25. It assesses the positive dimensions of behavior (under the category of adaptive behavior) and the negative (clinical) dimensions of behavior.

Designed to facilitate the diagnosis or educational classification of a variety of disorders or difficulties, the BASC-2 provides separate subscale scores for the clinical categories labeled aggression, anxiety, attention problems, atypicality, conduct problems, depression, hyperactivity, learning problems, somatization, and withdrawal; it also provides separate subscale scores for the adaptive functioning categories of adaptability, activities of daily living, functional communication, leadership, social skills, and study skills (Reynolds & Kamphaus, 2004). The various subscales are organized into Composites labeled Internalizing Problems, Externalizing Problems, Behavior Symptoms, School Problems and Adaptive Skills. Although the self-report forms essentially assess many similar aspects of behavior problems and adaptive functioning, items also address additional aspects and the item content is organized into different subscales and composites. The BASC-2 Self-Report Forms include the subscales labeled Attitude Toward School, Attitude Toward Teachers, Sensation Seeking, Atypicality, Locus Of Control, Social Stress, Anxiety, Depression, Sense of Inadequacy, Somatization, Hyperactivity, Attention Problems, Relationship with Parents, Interpersonal Relationships, Self-Esteem, and Self-Reliance. For each BASC-2 subscale and composite, t-scores (mean = 50; standard deviation of 10) and percentiles are provided along with information relative to the point at which scores are clinically significant, suggestive of a diagnosis or of a problem that would indicate treatment is needed, or suggest that the child or adolescent is at risk for a disorder.

Many of the specific items both of the rating scales and of the self-report that are included in various subscales describe behavior reflective of the use or disuse of self-regulation executive function capacities such as shifting, inhibiting, and sustaining. Specific examples include the Parent Form Aggression Subscale item “loses temper too easily;” the Attention Subscale items “pays inattention,” and “is easily distracted;” the Atypicality Subscale item “repeats thoughts over and over;” the Hyperactivity Subscale items ―cannot wait to take turn,” ―is unable to slow down,” ―disrupts other children‘s activities,” ―acts out of control,” ―is overly active,” ―acts without thinking,” and ―has poor self-control;” the Adaptability Subscale items ―adjusts well to changes in routine,” and ―adjusts well to changes in family plans;” the Conduct Problems Subscale items ―breaks the rules” and ―gets into trouble;” and the Leadership Subscale items ―is a self-starter,” ―makes decisions easily,” and ―gives good suggestions for solving problems.”

Relationship between Indirect and Direct Measures of Executive Functions

In the abundant literature on executive functions, little data can be found demonstrating that the deficits seen in test performance with direct measures are consistent with and related to the deficits assessed with the indirect measures, even though the language describing executive function deficits is similar in the manuals of direct and indirect assessment tools and the relationships are certainly implied in the findings. The BRIEF manual in particular contains language that strongly suggests that the BRIEF behavioral ratings for inhibition, shifting, emotional control, initiation, working memory (which is more suggestive of persistence and sustaining than more cognitively-based working memory skills), planning, organization of task approach and completion, as well as organization of materials, and monitoring of on-going activity assess the same executive functions as those measured by direct neuropsychological measures. Although the BRIEF manual discusses the scale‘s construct validity relative to other rating scales, the authors present no data relative to correlations between the rating item behavioral descriptions and neuropsychological test score data (Gioia et al., 2000).

However, in the discussion of the clinical scales, the authors do suggest a close relationship between behavior ratings and neuropsychological test performance.

There appears to be limited literature reporting correlations between ratings of behavior and performance on neuropsychological tests such as the WCST. However, given the increasing number of very recent journal articles on the topic of executive functions, this specific topic appears to be of growing interest. A study completed in 1994 (Riccio et al.) found no significant correlations of parent ratings on an Achenbach rating scale with WCST variables, although teacher ratings were more strongly correlated. Avila, Cuenca, Felix, Parcet, and Miranda (2004) assessed impulsivity in “nonselected school-aged boys” (n = 165) using different neuropsychological instruments and relating these results to parent and teacher ratings of hyperactivity and ODD. The goal of the study, according to its authors, was not only to assess whether or not there was a unitary factor to impulsivity, but also to link the neuropsychological measures with adult ratings of the boys. This study found a relationship between those neuropsychological measures which the authors identified as measures of behavioral inhibition (Stop Task, the Continuous Performance Task, the Matching Familiar Figures Test, and the Circle Tracing Task) and parent and teacher ratings of hyperactivity. However, there was not a significant relationship between WCST perseverative errors (considered to be the scoring variable most closely related to impulsivity) and parent and teacher ratings of hyperactivity. The authors indicated that the WCST was not a measure of inhibitory control, but of the ability to modify an ineffective response, suggesting that what was observed in impulsive behavior and what was measured on the WCST were two different executive deficits. A limitation of this study may be the fact that it was conducted with non-identified children, leading to questions relative to its ability to effectively characterize the executive function deficits more often evident in those with behavioral disorders.

This dearth of understanding regarding which, if any, of the neuropsychological instruments might have ecological validity is noted by Silver (2000) in her review of existing literature on this topic relative to children with Traumatic Brain Injury. Stating the degree to which neuropsychological test results can validly predict specific functioning in the natural environment…has received very little attention in the research literature" (p. 974), Silver further noted that even less attention has been given to whether or not there is consistency between parent ratings, from instruments such as the Vineland, and neuropsychological tests results. In noting the low to moderate correlations found in the existing literature between these assessment techniques, she echoes others in identifying the differing structure of the assessment tools as being a likely cause of the weaker correlations. However, she suggests that information compiled by both the more standardized and the narrative report assessments provide critical information about the impact of injury. The standardized assessments often identify those deficits that need to be addressed as well as the provide information relative to the conditions under which strengths are demonstrated, whereas the more narrative data, such as adaptive assessments and rating scales may better capture real-world functioning. Because her focus was on a population for which executive function deficits have been long reported, the point made by her is important: Data from multiple sources including neuropsychological test data, behavioral observations and rating scale results, as well as assessments of cognitive functioning, need to be integrated; the purposes are first, to inform understanding of strengths and weaknesses and the environments in which they are demonstrated and second, to direct intervention and treatment.

In their review, Alvarez and Emory (2006) echo the same theme, namely, that it is important to ground “the executive functions construct in the measurement of observable behaviors that have real-world significance” (p. 33). Progress toward this more comprehensive, real-world measurement may in fact come out of the rehabilitation literature, one article of which recommends both formal neuropsychological screening as well as use of caregiver rating scales, such as the BRIEF (Tarazi, Mahone, & Zabel, 2007). These authors also report on literature which suggests that task inhibition and set shifting difficulties, emanating from neuropsychological testing, were factors in differentiating those with TBI who needed self-care assistance from those that did not. This would suggest a correlation between test data and ecological functioning. Tarazi et al. go on to note that the World Health Organization, in its most recently proposed *International Classification of Functioning, Disability, and Health*, suggests that self-care capacities, mediated not only by the disability but also by the executive capacities of the individual, should dictate treatment and interventions.

In summary, an initial review of literature provides limited information relative to the degree to which there is uniformity in the constructs used to describe executive function difficulties in clinical assessments, both direct and indirect, of children and adolescents. This topic takes on particular importance in the practice of psychology in which there are often time-related or financial constraints on what measures can be utilized to assess presenting concerns. Of greater importance, because psychologists are called on to delineate interventions for these functional difficulties, it is important that they are able to be confident about the relationship between test data and constructs measured.

As the understanding of executive functions emerges and, specifically the understanding of the specific skills or deficits which are included in this more global term, it becomes increasingly important for clinicians to employ language relative to executive function capacities that is concise, descriptive and accurate. It is also important that clinicians use tools that characterize well the deficits noted. Studies that focus on understanding the specific executive function deficits that impede performance in multiple arenas serve to support the clinician's ability to communicate in this way. It is hoped that this study will increase that endeavor by examining the role and understanding of the relationship of specific executive function deficits evidenced in performance of the Wisconsin Card Sorting Test with other clinical measures as well as offer information relative to enhancing its use in school settings.

Statement of Problem

The Wisconsin Card Sorting Test remains one of the most frequently used evaluative tools in neuropsychological assessment, and has become even more frequently utilized with the increased focus on and understanding of what is encompassed under the executive function umbrella. In all likelihood because of the ambiguity of its directions and its multifaceted components, it is considered to be one of the best measures of executive function capacities, which, according to the global definition, is the ability to access and direct the use of cognitive abilities in the pursuit of a goal. The literature suggests that the WCST measures abstract reasoning, cognitive flexibility or the ability to shift set, working memory, and sustained attention; i.e, it measures many of those self-regulative executive capacities in the increasingly specific definition.

Difficulties in performance on the WCST have been well-documented in a variety of clinical populations. Historically, the instrument was used to identify frontal lobe injury in those with traumatic brain injury, Parkinson‘s disease, and schizophrenia, but more recently it has been used to identify frontal lobe deterioration in those with Alzheimer‘s and other dementias, as well as in those with a host of developmental or psychiatric issues. With the development of the Heaton version of 1993 and its standardization of children, it has been used increasingly to identify executive function difficulties in children in psychiatric populations.

However, in spite of its prevalent use and face validity, there remains considerable discussion about what the WCST measures. Although there appears to be relative consensus about the cluster of executive function and cognitive capacities that are required for adequate performance, there is far less consensus regarding the primary requirements for successful completion. Although considered a reasonable test of frontal or prefrontal lobe functioning, the literature tends to suggest that performance is not frontal lobe specific. Because most of the research has focused on adult psychiatric and other medical populations, there is even less clarity relative to the understanding of the factors that result in successful performance in children and adolescents, because there remains little literature about these probands.

Given the fact that many of the same self-regulation executive function capacities are the focus of other clinical instruments, such as behavior rating scales, like the BRIEF, other tests of executive functions, such as the D-KEFS, and even, to some extent, cognitive measures, such as the WISC-IV and WAIS-III, it appears important that the clinician understand degree to which there is similarity or difference between constructs being measured. To date, the literature appears somewhat inconclusive about whether or not the WCST measures of inhibition, set-shifting and working memory are similar to what is observed behaviorally as reflected in the scores derived from parent and teacher rating scales, such as the BRIEF, even though both types of measures are designed to assess executive functions and employ the same language to discuss executive function capacities. Additionally, the literature has provided little information regarding whether or not WCST scores are similar to the scores from measures of the D-KEFS, which provide numerous inhibition and switching scores. Finally, as the WCST is considered, to some extent, to be a measure of abstract reasoning, there is little information in the literature regarding the correlation between student performance on the WCST with measures of reasoning from the WISC-IV and WAIS-III.

Research Questions

Given the above discussion relative to some of the problems associated with use and interpretation of the WCST, this study will focus on answering three questions.

Question 1. The first question investigates the relationship between student performance on the Wisconsin Card Sorting Test and selected Subtest and Index scores from the Wechsler Scales, subtest scores measuring inhibition, shifting, and problem-solving capacities from the Delis-Kaplan Executive Function System and selected Scale and Subscale scores from the parent and teacher rating forms of the Behavior Rating Inventory of Executive Functions and the Behavior Assessment System for Children. Given the multifactorial nature of the WCST, the developmental variability of executive functions in children and adolescents, and the existing literature on the relationship of WCST results and other clinical measures, it is hypothesized that there will be a weak relationship between the WCST variables and these others measures.

Question 2. The second question investigates whether or not there are differences in student performance on the WCST by factors which include administration type, gender, diagnosis, and age groupings. Because the literature suggests some differences by age, it is hypothesized that some differences for this factor may be present, but not for the other factors of method of administration, diagnosis, or gender.

Question 3. Given the complexities of the constructs involved and the ambiguities with WCST interpretation, the third question of this study investigates whether or not a process-oriented approach to WCST performance provides important information that is not provided in the WCST scoring mechanisms. It is hoped that patterns and/or behavioral observations of performance will yield valuable information relative to assessment of executive functions, using the WCST.

Chapter Two

Methodology

The purpose of this study was to investigate the relationship between student performance on the Wisconsin Card Sorting Test and other clinical measures, including selected tasks from the D-KEFS, the Wechsler Scales, the BRIEF, and the BASC, instruments which purport to measure constructs similar to those thought to be assessed by the WCST.

Participants

The sample for this study consisted of data from the archival records of 94 children from the ages of 8 through 19 who underwent psychological assessments between 2000 and 2008. These archival records were collected on children and adolescents from Connecticut, New York, and Pennsylvania; these children were referred either to their schools' evaluative teams or were referred by their parents for private evaluation; these evaluations were completed by four different school psychologists. Of the 94 students, 65 were male and 29 were female. The students demonstrated specific deficits in behavioral or emotional regulation that had been identified either prior to or as a result of the evaluative process. Ten students had an IDEA eligibility of ADHD; 22, of Specific Learning Disability; 3, Emotional Disturbance; 3, Speech and Language Impairment; 1, Pervasive Developmental Disorder and 1, Hearing Impairment; the remaining 54 had not yet been classified because of their initial referral status or did not meet eligibility requirements for IDEA classification. Thirty-three students had diagnoses of ADHD (all variants of that diagnosis) alone or ADHD with Bipolar Disorder, OCD, ODD, depression, Tic Disorder, and Generalized Anxiety Disorder; two carried the diagnosis of Bipolar Disorder only; 5, depression only; and 4, anxiety disorders. No data were collected for the purposes of this study relative to whether or not the students were receiving pharmacological treatment.

In most cases these students were experiencing academic difficulties; however, all but three had earned Full Scale IQ scores of 85 or higher. The remaining three had earned Full Scale IQ scores below 85, but had earned either a Perceptual Reasoning Index or Verbal Comprehension Index score that fell within the average range. Data obtained from the records of these children included score data from the following assessments: the Behavior Rating Inventory of Executive Functions (BRIEF), the Behavior Assessment System for Children (BASC or BASC-2); the Wechsler Scales (WISC-IV or WAIS-IV), the Wisconsin Card Sorting Test (WCST), and the Delis-Kaplan Executive Function System (D-KEFS). Of the 94 students, 64 had been administered the manual version of the WCST and 30 had been administered the computerized version.

Variables

Variables included in the study came from a number of different measures.

WCST. Variable obtained from the WCST included the following standard scores (mean 100, standard deviation 15): Percent Total Number Correct, Percent Errors, Percent Perseverative Responses, Percent Perseverative Errors, Percent Nonperseverative Errors, Percent Conceptual Level Responses, and the following raw scores: number of Categories Completed, and number of trials to complete each category. The Heaton manual (Heaton et al., 1993) provides data relative to the WCST's reliability: Interscorer and intrascorer reliabilities were deemed to be excellent (reliability coefficients ranging from .83 to 1.00 for all of the scores used in this study) on WCST data both for children and for adult populations. Using a generalizability theory model, data suggested strong reliability for all variables except Percent Perseverative Responses and Percent Perseverative Errors, which showed moderate reliability. Standard errors of measurement ranged from 7.94 to 11.91 (Mean = 100, standard deviation = 15) in a sample of children and adolescents. As noted previously, given the WCST's extensive use in clinical and research populations, Heaton et al. note its validity as a measure of executive functions.

Wechsler Scales. The data set used in this study included scores from the WAIS-III (n = 25) and the WISC-IV (n = 60). The variables from the WISC-IV utilized in this study included the Matrix Reasoning Subtest scaled score, Full Scale IQ score, and Verbal Comprehension, Perceptual Reasoning, Working Memory and Processing Speed Index scores. From the WAIS-III, the same variables were examined, with full recognition of the subtest composition difference between the WAIS-III Perceptual Organization and WISC-IV Perceptual Reasoning Indexes and between the WAIS-III and WISC-IV Verbal Comprehension Indexes. The WISC-IV was standardized from a stratified sample of 2,200 children as well as samples from special groups, ages 6.0 to 16.11 years (Wechsler, 2004). The WAIS-III standardization sample totaled 2,450 individuals, ages 16-89 years. Internal reliability coefficients for WISC-IV measures used were high, with the reliability coefficient of .97 for Full Scale score; a range of .88

(Processing Speed) to .94 (Verbal Comprehension) for the composites; and a .89 for Matrix Reasoning. Test-retest stability coefficients were only slightly lower, with a range of .85 to .93 for these same WISC-IV variables. Standard error of measurement for Matrix Reasoning was a .99 (mean = 10, standard deviation = 3); standard errors of measurement for the composite and Full Scale scores ranged from 2.68 (Full Scale) to 5.21 (Processing Speed) (mean = 100; standard deviation =15). The WISC-IV technical manual (Wechsler) provides extensive information relative to the test‘s validity, citing exploratory and confirmatory factor-analytic studies and correlational studies with other measures and in special populations. Reliability coefficients for WAIS-III Matrix Reasoning for the 16-19 age groups were .87 and .89; coefficients for the composite and Full Scale variables for these age groups ranged from .86 (Processing Speed) to .97 (Full Scale and Verbal Comprehension). Standard error of measurement for Matrix Reasoning for the 16-17 age group was 1.08 and for 18-19 group, .99; for the composite and Full Scale scores the range of the standard errors of measurement was 2.38 to 5.56. Validity data and results similar to the WISC-IV data are found in the technical manual.

D-KEFS. Variables from the D-KEFS used in the study were the Inhibition and Inhibition/Switching scaled scores for the Color Word Interference Subtest, and the Total Questions Asked scaled score of the Twenty Questions Subtest. Although the D-KEFS Sorting Test is thought to assess cognitive capacities similar to those assessed by the WCST, it was usually excluded from the assessments administered to the subjects of this study because of the amount of time required for administration and the anticipated overlap with the score information gathered with the WCST. The D-KEFS was standardized on a stratified sample of 1750 individuals ranging in age of 8-89 years. In the technical manual, Delis, Kaplan, and Kramer (2001) note the complexities in providing reliability data for the individual subtests, because of multiple factors in each subtest. Thus reliability coefficients are highly variable, depending on the age group, task, and psychometric property. Test-retest reliability coefficient for the age group of this data set (8-19) for Color-Word Inhibition was .90, and for Color-Word Inhibition/Switching, .80. Standard errors of measurement for this age group ranged from 1.38 to 1.85. The reliability coefficient was not available for Tower Achievement Total Questions Asked, although for Total Weighted Achievement, a similar measure, the coefficient was quite low (.06) for this age group, with standard errors of measurement ranging from 2.11 to 2.85. Internal reliability coefficient for Total Weighted Achievement for ages 8-19 was .10 to .53. As noted previously, D-KEFS validity studies are few, because of the test's relatively young age; the authors note the test's validity based on its similarities to older tasks for which there is considerable validity data.

BRIEF. Variables from the BRIEF used in this study include the T-scores from the Inhibit, Shift, Working Memory, and Monitor Scales. The BRIEF was standardized with 1,419 parent and 720 teacher ratings of children, ages 5-18 years, in a normative sample designed to reflect U. S. demographic groups. Internal consistency coefficients for the parent form for Inhibition was .94 (clinical sample) and .91 (normative sample); Shift, .88 and .81; Working Memory, .92 and .93; and Monitor, .85 and. 83. For the teacher form, internal consistency coefficients were Inhibition, .95 and .96; Shift, .91 (for both samples); Working Memory, .90 and .93; and Monitor, .89 and .90. Test-retest reliability for teacher ratings for these four scales ranged from .83 to .91; the range was greater for parent ratings (both normative and clinical populations) for the four scales, from .72 to .84. It is important to note that correlations within and between parent and teacher ratings (for the normative sample) range from moderate to low: For the Inhibit scale, .50; Shift scale, .15; Working Memory, .30; and Monitor scale, .42. Analyses of correlation with other behavior rating scales suggested convergent and divergent validity with similar construct measures, but less strong correlations with broader measures of emotional functioning (Gioia et al., 2000).

BASC and BASC-II. Variables from the BASC and BASC-II used in this study included the Attention Problems, Conduct Problems, Aggression and Leadership Subscale T-scores obtained with the Parent and Teacher rating forms. The BASC and BASC-2 manuals provide extensive information relative to standardization procedures, reliability tests of internal consistency, test-retest reliability, interrater reliability, as well as validity data including factor analyses and correlations with other clinical measures (Reynolds & Kamphaus; 1993; 2004). BASC-2 Teacher ratings reliability coefficients (coefficient alpha value) for the age ranges included in this study for Attention Problems range from .89 to .95; for Conduct Problems, .90 to .92; for Aggression, .88 to .93; and for Leadership, .85 to .88. Parent ratings reliability coefficients (coefficient alpha values) for these ages for Attention Problems were .85 to .90; for Conduct Problems, .83 to .88; for Aggression, .84 to .89; and for Leadership, .84 to .86. Teacher ratings test-retest reliabilities for each scale ranged from .78 to .90; on the parent ratings for these scales, the range was .72 to .87. Correlations between teacher and parent ratings for these scores for combined general and clinical samples, however, are much lower, with coefficient of .38 for aggression on the child form and .44 on the adolescent form; for conduct problems, .43 on the child form and .49, adolescent; for attention problems, .52, child, and .46, adolescent; and for leadership, .43, child, and .39, adolescent. BASC Teacher ratings reliability coefficients (coefficient alpha values) for the age ranges included in this study for Attention Problems range from .86 to .93; for Conduct Problems, .67 to .92; for Aggression, .92 to .95; and for Leadership, .88 to .92. Parent ratings reliability coefficients (coefficient alpha values) for these ages for Attention Problems were .79 to .83; for Conduct Problems, .64 to .75; for Aggression, .77 to .84; and for Leadership, .82 to .88. Teacher ratings test-retest reliabilities for each scale ranged from .78 to .93; on the parent ratings for these scales the range was .58 to .92. Correlations between teacher and parent ratings for these scores for combined general and clinical samples, however, are much lower, with a coefficient of .38 for aggression on the child form and .35 on the adolescent form; for conduct problems, .49 on the child form and .63, adolesecent; for attention problems, .62 on the child form and .48, adolescent; and leadership, .52 on the child form, and .47 on the adolescent form.

Overview of Research Design

A descriptive and correlational research design was used for this study. Correlational and cross-tabulation analyses and parametric inferential statistical tests were conducted using the data set compiled from the archival records. Where sample size allowed, analyses compared subgroups based on differences in form of WCST administration, gender, diagnostic categories, and age groupings. Descriptive data was provided, detailing patterns of performance on the WCST not reflected in WCST standard scores. Data analyses were conducted using descriptive frequencies, correlations, crosstabulations, and analyses of variance.

Chapter Three

Results

Results of this study are presented in three parts corresponding to the three research questions. The results related to the first question summarize descriptive statistics, correlations, cross-tabulation analyses and ANOVA‘s, utilizing the WCST and other clinical measure variables. The results related to the second question involved cross-tabulation analyses and ANOVA‘s to assess differences among subgroups. The results related to the third question consisted of descriptive statistics, cross-tabulation analyses, and observational data relative to patterns of performance, specific cases, or groups of cases.

Question 1: What is the relationship between student performance on the WCST and performance on selected Subtests and Indexes of the Wechsler Scales, selected D-KEFS measures of inhibition, shifting, and problem-solving; and parent and teacher ratings on the BRIEF Inhibit, Shift, Working Memory, and Monitor scales and parent and teacher ratings of BASC and BASC-2 scales of Aggression, Conduct Problems, Attention Problems, and Leadership?

Tables 1 and 2 profile the means and standard deviations for the variables selected from the tests administered to the 94 students. It is important to note that not all subjects were administered all of the measures used in this study, resulting in *n* counts of varying size for many of the variables listed in Tables 1 and 2.

Table 1

Descriptive Statistics for Executive Function and Cognitive Measures

Variable	*n*	*M*	*SD*
WCST Error %[1]	94	104.89	16.50
WCST PersResp%[1]	72	112.24	24.84
WCST PersError%[1]	93	113.68	23.86
WCST NonpersError%[1]	93	99.12	14.08
WCST ConLRes%[1]	94	107.07	19.15
WCST CatCom	94	5.22	1.40
Wec FS[1]	79	107.97	12.86
Wec PR[1]	82	108.84	12.84
Wec VC[1]	82	112.85	15.14
Wec PS[1]	81	95.59	13.02
Wec WM[1]	81	100.53	14.49
Wec MR[2]	84	11.99	2.72
D-KEFS CWI[2]	64	9.56	3.00
D-KEFS CWI/S[2]	63	9.52	3.44
D-KEFS TQTQA[2]	68	10.25	2.68

Note. WCST PersResp = Percent Perseverative Responses; WCST PersError = Percent Perseverative Errors; WCST NonpersError = Percent Non-Perseverative Errors; WCST ConLRes = Percent Conceptual Level Responses; WCST CatCom = Categories Completed; Wec FS = Wechsler Full Scale; Wec PR = Perceptual Reasoning, Wec VC = Verbal Comprehension; Wec PS = Processing Speed; Wec WM = Working Memory; Wec MR = Matrix Reasoning; D-KEFS CWI = Color-Word Inhibition; D-KEFS CWI/S = Color-Word Inhibition/Switching; D-KEFS TQTQZA = Twenty Questions Total Questions Asked.
[1]Standard Score (*M* = 100; *SD* = 15)
[2]Scaled Score (*M* = 10; *SD* = 3)

Table 2

Descriptive Statistics for Behavior Rating Scale Measures

Variable	*n*	*M*	*SD*
BRIEF P-In	75	57.63	13.08
BRIEF P-Sh	75	59.97	13.08
BRIEF P-WM	75	65.47	11.81
BRIEF P-M	75	62.96	11.66
BRIEF T-In	48	58.44	13.27
BRIEF T-Sh	48	64.42	16.53
BRIEF T-WM	48	70.88	13.40
BRIEF T-M	48	66.56	12.08
BASC P-CP	66	54.24	12.51
BASC P-Agg	66	51.80	10.92
BASC P-Att	66	62.74	9.39
BASC P-Lead	66	44.44	9.86
BASC T-CP	49	52.59	10.55
BASC T-Agg	49	51.61	10.47
BASC T-Att	49	60.41	9.28
BASC T-Lead	49	41.67	7.59

Note. BRIEF P-In = Parent Inhibit Scale; BRIEF P-Sh = Parent Shift Scale; BRIEF P-WM = Parent Working Memory Scale; BRIEF P-M = Parent Monitor Scale; BRIEF T-In = Teacher Inhibit Scale; BRIEF T-Sh = Teacher Shift Scale; BRIEF T-WM = Teacher Working Memory Scale; BRIEF T-M = Teacher Monitor Scale; BASC P-CP = Parent Conduct Problems Scale; BASC P-Agg = Parent Aggression Scale; BASC P-Att = Parent Attention Problems Scale; BASC P-Lead = Parent Leadership Scale; BASC TCP = Teacher Conduct Problems Scale; BASC T-Agg = Teacher Aggression Scale; BASC T-Att = Teacher Attention Problems Scale; BASC T-Lead = Teacher Leadership Scale.
T-Score (M = 50; SD = 5)

WCST Error Percentage, Non-Perseverative Error Percentatge and Conceptual Level Response Standard Score means fell into the average range, but Perseverative Response and Perseverative Error mean scores were in the high average range. Wechsler

Scale Index Standard Score and Subtest Scaled Score means were in the average range with the exception of the Verbal Comprehension Index Standard score mean, which was in the high average range. D-KEF's scores were in the average range. BRIEF ratings reflected means in the clinically significant range for parent and teacher form Working Memory Scale T-scores and for teacher form Monitor Scale T-scores. BASC ratings reflected a mean score in the clinically significant range for the parent and teacher form Attention Problems Subscale.

Although each of these test score variables (with the exception of the WCST Categories Completed raw score) are thought to be normally distributed in the general population, the assumption of linearity was not clearly met; therefore, Spearman's rho correlations were computed to examine the relationships among the WCST, Wechsler, D-KEFS, BASC and BRIEF variables. Table 3 shows the correlations among scores from the WCST and the other clinical measures used in the study. Two of the five WCST indices (Error Percentage and Conceptual Level Response Percentage) were significantly ($p < .05$) but very modestly correlated with Wechsler Full Scale, Perceptual Reasoning, Verbal Comprehension, and Processing Speed Index scores and the Matrix Reasoning Subtest score, but not significantly correlated with the Working Memory Index score. Two other WCST indices (Perseverative Response Percentage and Perseverative Error Percentage) were significantly ($p < .05$) but very modestly correlated with the Wechsler Full Scale and Matrix Reasoning Subtest scores, with the Perseverative Error Percentage

Table 3

Correlations Between WCST Variables and Other Clinical Measures

Clinical Variable	WCST Variable Error	PerResp	PerErr	NPerErr	ConLResp	CatCom
Wec FS	.36**	.31*	.31**	.13	.36**	.15
Wec PR	.35**	.24	.21	.25*	.30**	.10
Wec VC	.33**	.23	.22	.21	.35**	.17
Wec PS	.30**	.24	.26*	.13	.32**	.22*
Wec WM	.13	.09	.14	.04	.11	.15
Wec MR	.33**	.32**	.29**	.15	.27*	.20
D-KEFS CWI	.19	.19	.20	.09	.17	.18
D-KEFS CWI/S	.18	.13	.17	.06	.14	.15
D-KEFS TQTQA	.30*	.33*	.35**	.16	.34**	.28*
BRIEF P-In	-.24*	-.38**	-.29*	.04	-.29*	-.25*
BRIEF P-Sh	-.35**	-.14	-.18	-.18	-.34**	-.26*
BRIEF P-WM	-.09	-.04	-.01	-.09	-.16	-.14
BRIEF P-M	-.05	-.04	-.04	.05	-.10	-.14

Note. WCST PerResp = Percent Perseverative Responses; WCST PerErr = Percent Perseverative Errors; WCST NPerErr = Percent NonPerseverative Error; WCST ConLRes = Percent Conceptual Level Responses; WCST CatCom = Categories Completed; Wec FS = Wechsler Full Scale; Wec PR = Perceptual Reasoning, Wec VC = Verbal Comprehension; Wec PS = Processing Speed; Wec WM = Working Memory; Wec MR = Matrix Reasoning; D-KEFS CWI = Color Word Interference Inhibition; D-KEFS CWI/S = Color Word Interference Inhibition/Switching; D-KEFS TQTQZA = Twenty Questions Total Questions Asked; BRIEF P-In = Parent Inhibit Scale; BRIEF P-Sh = Parent Shift Scale; BRIEF P-WM = Parent Working Memory Scale; BRIEF P-M = Parent Monitor Scale
** $p < .01$; * $p < .05$

Table 3 (cont.)

Correlations Between WCST Variables and Other Clinical Measures

Clinical Variable	Error	PerResp	PerErr	NPerErr	ConLResp	CatCom
BRIEF T-In	-.23	-.43*	-.35*	.08	-.26	-.19
BRIEF T-Sh	-.19	-.10	-.15	.10	-.21	-.09
BRIEF T-WM	-.29*	-.44*	-.31*	-.09	-.32*	-.26
BRIEF T-M	-.07	-.21	-.18	.26	-.11	-.13
BASC P-CP	.02	.08	.04	-.06	.05	.15
BASC P-Agg	-.02	.11	.09	-.08	-.02	.17
BASC P-Att	-.23	-.21	-.23	-.07	-.24	-.12
BASC P-Lead	.12	.08	.08	.05	.18	.09
BASC T-CP	-.14	.01	-.06	-.06	-.09	.04
BASC T-Agg	-.23	-.12	-.16	-.07	-.25	-.13
BASC T-Att	-.22	-.14	-.18	-.00	-.22	-.05
BASC T-Lead	-.04	-.10	-.08	-.09	.00	-.07

Note. BRIEF T-In = Teacher Inhibit Scale; BRIEF T-Sh = Teacher Shift Scale; BRIEF T-WM = Teacher Working Memory Scale; BRIEF T-M = Teacher Monitor Scale. BASC P-CP = Parent Conduct Problems Scale; BASC P-Agg = Parent Aggression Scale; BASC P-Att = Parent Attention Problems Scale; BASC P-Lead = Parent Leadership Scale; BASC TCP = Teacher Conduct Problems Scale; BASC T-Agg = Teacher Aggression Scale; BASC T-Att = Teacher Attention Problems Scale; BASC T-Lead = Teacher Leadership Scale.
** $p < .01$
* $p < .05$

score also significantly and modestly correlated with Wechsler Processing Speed Index score. The WCST Nonperseverative Errors Percentage score was significantly but very modestly correlated only with the Wechsler Perceptual Reasoning Index score. The WCST Categories Completed score was significantly but very modestly correlated only with Wechsler Processing Speed Index score. Calculation of r^2 for each of these significant correlations indicates that the shared variance between WCST scores and Wechsler Full Scale, Verbal Comprehension, Perceptual Reasoning, Processing Speed Index, or Matrix Reasoning Subtest scores is less than thirteen percent in all cases and less than five percent in a number of specific instances.

With the exception of the Nonperseverative Errors Percentage score, WCST indices scores were significantly ($p < .05$) but very modestly correlated only with the D-KEFS Total Questions Asked score; none of the WCST indices were significantly correlated with the D-KEFS Inhibition or Inhibition/Switching scores.

All of the WCST variables except Nonperseverative Errors Percentage were significantly ($p < .05$) but very modestly correlated with the BRIEF Parent form Inhibition Scale and the BRIEF Teacher form Working Memory Scale. Additionally, the WCST Error Percentage, Conceptual Level Response Percentage, and Categories Completed scores were significantly ($p < .05$) but very modestly correlated with the BRIEF Parent Inhibit Scale score, and the WCST Perseverative Response Percentage and Error Response Percentage were significantly ($p < .05$) but modestly correlated with the BRIEF Teacher form Inhibit Scale scores. As with the other variable relationships, the large majority of these correlations indicate there is less than ten percent of shared variance in WCST scores and BRIEF Inhibit, Working Memory and/or Shift Scale scores obtained from parent and teacher ratings.

None of the WCST indices were significantly correlated with the BRIEF Parent form Working Memory or Monitor Scale scores or with the BRIEF Teacher form Shift and Monitor Scale scores. Similarly, none of the BASC Subscales used in this study (Conduct Problems, Aggression, Attention Problems, and Leadership) were significantly correlated with any of the WCST scores.

In terms of clinical relevance, the correlations between WCST score indices and select Wechsler, D-KEFS, BRIEF, and BASC scores are much lower than would be expected if the WCST scores reflected performance of constructs similar to those assessed with these other clinical instruments. This is especially true in the case when compared with the data typically reported from the intercorrelations of cognitive assessment instruments thought to be measuring similar cognitive capacities. Given the modest levels of correlation between WCST indices and the scores from other clinical measures, these results suggest little in the way of a meaningful relationship between performance on the WCST and other cognitive, executive function, and behavioral measures thought to be measuring similar or related cognitive constructs.

A cross-tabulation of WCST Error Percentage and Conceptual Level Response Percentage Scores by Wechsler score categories highlights the wide variability of the relationship between the two measures. Standard scores were recoded into categorical variables (scores less than 70 = 1; 71-89 = 2; 90-99 = 3; 100-109 = 4; 110-119 = 5;

Table 4

Cross-Tabulation of WCST and Other Variables by Score Groupings

Variable	WCST Score Range in Relation to Score Ranges on Other Variables		
Comparison	Lower	Comparable	Higher
WCST Error Percentage:			
Wec FS (*n* = 79)	27 (34%)	21 (27%)	31 (39%)
Wec PR (n = 82)	35 (43%)	16 (19%)	31 (38%)
Wec VC (n = 82)	38 (46%)	26 (32%)	18 (22%)
Wec PS (n = 81)	16 (20%)	14 (17%)	51 (63%)
Wec WM (n = 81)	21 (26%)	16 (20%)	44 (54%)
Wec MR (n = 84)	38 (45%)	26 (31%)	20 (24%)
D-KEFS CWI (n = 64)	18 (28%)	12 (19%)	34 (53%)
D-KEFS CWI/S (n = 63)	19 (30%)	12 (19%)	32 (51%)
D-KEFS TQTQA (n = 68)	23 (34%)	16 (23%)	29 (43%)
WCST Conceptual Level Response Percentage:			
Wec FS (n = 79)	26 (33%)	23 (29%)	30 (38%)
Wec PR (n = 82)	32 (39%)	20 (24%)	30 (37%)
Wec VC (n = 82)	34 (41%)	31 (38%)	17 (21%)
Wec PS (n = 81)	16 (20%)	11 (13%)	54 (67%)
Wec WM (n = 81)	19 (23%)	20 (25%)	42 (52%)
Wec MR (n = 84)	38 (45%)	25 (30%)	21 (25%)
D-KEFS CWI (n = 64)	17 (26%)	12 (19%)	35 (55%)
D-KEFS CWI/S (n = 63)	17 (27%)	14 (22%)	32 (51%)
D-KEFS TQTQA (n = 68)	19 (28%)	18 (26%)	31 (46%)

Note. Wec FS = Wechsler Full Scale; Wec PR = Perceptual Reasoning, Wec VC = Verbal Comprehension; Wec PS = Processing Speed; Wec WM = Working Memory; Wec MR = Matrix Reasoning; D-KEFS CWI = Color-Word Inhibition; D-KEFS CWI/S = Color-Word Inhibition/Switching; D-KEFS TQTQZA = Twenty Questions Total Questions Asked.

greater than or equal to 120 = 6). As noted in Table 4, cross-tabulations revealed a relatively even distribution of Full Scale and Perceptual Reasoning Index score cateories across the categories both of WCST Error Percentage and of Conceptual Level Response Percentage, with almost equal numbers scoring higher or lower on the Wechsler variables compared with the WCST variables. More uneven distributions are found when the other Wechsler Index scores are compared with these WCST variables. More students of this referred population tended to earn better scores on the WCST Error Percentage and Conceptual Level Percentage variables than on the Wechsler Working Memory and Processing Speed Indexes. However, cross-tabulations indicated that a greater number of students earned lower scores on the WCST variables than on the Wechsler Verbal Comprehension Index and the Matrix Reasoning Subtest. Although the distribution of scores within categories was not quite as skewed, a greater number of students of this sample also tended to earn higher scores on the two WCST variables than on the D-KEFS variables.

Because the literature suggests a strong association between cognitive abilities, especially reasoning, and performance on the WCST, this study examined more closely the distribution of scores for WCST Conceptual Level Responses and Wechsler Full Scale and Matrix Reasoning variables. Table 5 shows the distribution of Conceptual Level Responses by Wechsler Full Scale scores, utilizing the categorical variables previously delineated. In the highest category (>119), most (64%) of the students in the data set achieved comparable Full Scale scores. However, for those scoring 119 or lower

Table 5

WCST Percent Conceptual Level Response Range by Wechsler FSIQ Score Range

FSIQ Range	WCST Score Range		
	Lower	Comparable	Higher
> 119	5 (36%)	**9 (64%)**	0 (0%)
110-119	8 (50%)	**4 (25%)**	4 (25%)
100-109	8 (26%)	**8 (26%)**	15 (48%)
90-99	4 (36%)	**1 (7%)**	8 (57%)
71-89	0 (0%)	**0 (0%)**	3 (100%)

on Conceptual Level Responses, the distribution is much more dispersed, suggesting that the Wechsler Full Sclae score—with its composition of multiple components—is not likely to be helpful in understanding performance on the WCST or, at least, the self-regulation of reasoning that is suggested by the Conceptual Level Response score.

When the relationship between Wechsler Matrix Reasoning scores and WCST Conceptual Level Response scores is examined, a more interesting result is found that suggests to what extent reasoning capacities may, in fact, be related to performance of the WCST as measured by Conceptual Level Responses. Table 6 provides the distribution of scores by categorical variables for Wechsler Matrix Reasoning and WCST Conceptual Level Response. Of those 50 students who achieved a score of 12 or greater (recomputed as a Standard Score of 110) on Matrix Reasoning, only 4 earned a score of 110 or higher on WCST Conceptual Level Responses. Thus, ninety-two percent of the students achieved Matrix Reasoning scores that were higher than their WCST scores. Although WCST results were more scattered for individuals who achieved Matrix Reasoning

Table 6

WCST Percent Conceptual Level Response Range by Wechsler Matrix Reasoning Score Range

Matrix Reasoning Range	WCST Score Range		
	Lower	Comparable	Higher
> 119	16 (59%)	**11 (41%)**	0 (0%)
110-119	11 (48%)	**8 (35%)**	4 (17%)
100-109	9 (43%)	**5 (24%)**	7 (33%)
90-99	0 (0%)	**1 (17%)**	5 (83%)
71-89	2 (29%)	**0 (0%)**	5 (71%)

scores below 12, with those scoring below ten showing stronger relative Conceptual Level Response scores, overall results found that seventy-five percent of the students performed better on Matrix Reasoning than on Conceptual Level Responses. For the students in this study, then, it was difficult to earn higher WCST Conceptual Level Response scores than Matrix Reasoning scores, especially at the higher cognitive levels. When these results are viewed in light of the theoretical connection between reasoning ability and self-regulation executive cueing of reasoning ability, these results, as will be discussed later, suggest that Matrix Reasoning may represent a baseline of abstract reasoning capacities, whereas the Conceptual Level Responses score reflects the self-regulation executive control of those reasoning capacities.

These categorical variables, by definition, allowed for only an *n* of 0 for scoring higher on WCST Conceptual Level responses than on Matrix Reasoning. In fact,

Table 7

WCST Percent Conceptual Level Response Range by Wechsler Processing Speed Score Range

Processing Speed Range	WCST Score Range		
	Lower	Comparable	Higher
> 119	2 (100%)	**0 (0%)**	0 (0%)
110-119	4 (33%)	**4 (33%)**	4 (33%)
100-109	5 (24%)	**3 (14%)**	13 (62%)
90-99	9 (31%)	**2 (12%)**	18 (62%)
71-89	0 (0%)	**5 (25%)**	15 (75%)
< 71	0 (0%)	**0 (0%)**	6 (100%)

however, in that group of 11 who received comparable scores in the greater than 120 range, there were four individuals who did significantly better (equal to or greater than 1 standard deviation) on Conceptual Level Responses than on Matrix Reasoning. In looking at the four of this group and the four in the next group who earned Conceptual Level Response scores higher than Matrix Reasoning scores, all but two had Processing Speed scores that were significantly below their WCST scores. The range of Conceptual Level Response scores was 120-146 (with 6 receiving a score of 146), with a mean score of 139; however, the Processing Speed scores ranged from 84 to 117, with a mean of 96. Table 7 shows the distribution of Processing Speed categorical scores by Conceptual Level Response. A majority of students who scored average or well below average on the Wechsler Processing Speed tasks scored above that range on Conceptual Level Responses, suggesting that slower processing appears to support WCST performance at the higher performance levels. Because this discrepancy is not maintained at the lower

WCST ranges, there is the suggestion that slower processing speed may be a moderating factor in WCST Conceptual Level performance only at higher score ranges. What is also interesting about this group of eight students is that all but two were 14 years of age or older; the other two were nine and twelve. As will be discussed in the next section, at the lower age range (8-11) there is some suggestion that faster rather than slower processing speed is a positive factor in WCST performance.

Question 2: Are there differences in student performance by factors including administration type, gender, diagnosis and age groupings?

The second question of this study focused on ipsative data from the WCST, as well as relationships of variables according to sample subgroup factors. Although the number of male students in the data set was more than two times the number of female students, the overall size of each subgroup allowed an investigation of performance by gender.

Gender. Gender comparisons were statistically tested using one-way ANOVA's. No statistical differences were found between subjects, based on gender for the variables WCST Error Percentage (F (1,92) = .33, p = .56, η^2 = .004); Perseverative Responses Percentage (F (1,70) = .12, p = .73, η^2 = .002); Perseverative Error Percentage (F (1,91) = .01, p = .92, η^2 = .00); Non-Perseverative Error Percentage (F (1,91) = .84, p = .36, η^2 = .01); Conceptual Level Responses Percentage (F (1,92) = .59, p = .44, η^2 = .01); and Categories Completed (F (1,92) = .15, p = .69, η^2 = .002).

Administration Method. ANOVA tests conducted on method of administration suggest that performance on the WCST may be somewhat influenced by whether or not the student was administered the computer version or the manual version of the test. There were no significant differences between computer-administration and manual-administration groups for Error Percentage (F (1, 92) = .43, p = .51, η^2 = .005); Perseverative Response Percentage (F (1,70) = 2.08, = .15, η^2 = .03); Conceptual Level Response Percentage (F (1,92) = .14, p = .71, η^2 = .001); and Categories Completed (F (1,92) = .13, p = .72, η^2 = .001). However, there were significant differences between computer-administration and manual-administration groups for Perseverative Error Percentage (F (1,92) = 4.56, p = .03, η^2 = .05) and Non-Perseverative Error Percentage (F (1,91) = 9.17, p = .003, η^2 = .09). The Nonperseverative Errors mean score of the WCST computer-administration group (mean = 105.41) was greater than the mean score of the WCST manual-administration group (mean = 96.27), but the computer-administration group's Perseverative Errors Percentage mean score (mean = 106.17) was lower than the manual-administration group's mean score (mean = 117.25).

Cross-tabulation of WCST score ranges (using the scores categories described in the previous section) by administration method indicate that ninety-one percent (32 of 35) of those who obtained a WCST Perseverative Error score of 120 or higher were in the manual administration group, with less significant or no differences by administration were observed in the lower than 120 groups. Thus, the difference in Perseverative Error scores by administration format may be related to a sampling artifact of this particular data set.

Clinical Diagnoses. One-way ANOVAs did not distinguish the ADHD probands from those with other diagnoses or with no known diagnosis in this referred population.

When the ADHD probands were clustered with those with the diagnosis of Bipolar Disorder, there was a trend toward a significant difference in Non-Perseverative Response Percentage (F (1,91) = 3.78, p = .05, η^2 = .01), suggesting that subjects displaying more dysregulated behaviors may be more at-risk for random responding than those with other or no diagnoses. However, given the fact that all the students referred for evaluation were experiencing some level of academic, behavioral, and/or emotional difficulties, these results may need to be interpreted with caution, because those in the other-than-ADHD group may have been diagnosed with ADHD or Bipolar disorder at some point in time after the evaluation data had been collected.

Chronological Age. To assess whether or not age might have been a factor in WCST performance, the 94 cases were clustered by age groupings of 8-11 years, 12-15 years, and 16-19 years and one-way ANOVA's were conducted to determine whether or not there were significant differences in performance on the WCST variables targeted. A statistically significant difference was found among the three age groupings for Error Percentage (F (2,91) = 4.35, p = .02, η^2 = .09); Perseverative Response Percentage (F (2,69) = 4.87, p = .01, η^2 = .12); Perseverative Error Percentage (F (2,90) = 5.00, p = .01, η^2 = .10); Conceptual Level Responses (F (2,91) = 6.02, p = .003, η^2 = .12); and Categories Completed (F (2,91) = 6.17, p = .003, η = .12), but not Non-Perseverative Error Percentage (F (2,90) = 1.48, p = .23, η^2 = .03). Post hoc Scheffe Tests indicate that students in the 16-19 age group performed significantly better than the 8-11 age group on Total Errors, Perseverative Responses, Perseverative Errors, and Conceptual Level Responses. The middle group outperformed the youngest group on Perseverative Error

Percentage. Because assumption of equality of variance was not met in the case of the Categories Completed score, an alternate post hoc significance test was applied (Games-Howell test), indicating that both the middle and the oldest group completed significantly more categories than the youngest group. These results are interesting in light of the fact that these scores, with the exception of Categories Completed, are age-corrected standard scores. When these three age groups were compared relative to the other variables of the data set, the only differences found occurred with the Matrix Reasoning score (F (2, 81) = 4.03, p = .02, η =.09), with the oldest group outperforming the youngest, and the D-KEFS Twenty Questions Total Questions Asked score (F (2, 65) = 3.69, p = .03, η = .03), with the students in the 12-15 age group outperforming the youngest group. Variance tests for homogeneity were not met for Wechsler Full Scale, Perceptual Reasoning, and Matrix Reasoning scores.

When correlations were computed for this youngest age group, WCST Error Percentage scores and WCST Conceptual Level Percentage scores were weakly, but *negatively* correlated with every Wechsler score, although not at a level of significance. The only exception was the correlation with Wechsler Processing Speed Index scores, in which the correlations were .41 and .43 (p < .05). Perseverative Error percentage scores had a stronger correlation (r =.48, p < .05) with Processing Speed Index scores as well. These results may be an artifact of this particular sample; however, it may suggest that faster processing speed capacities might be more impactful in WCST performance at younger ages. (As previously noted, for the older age group (16-19) there appears to be an inverse relationship, because lower Processing Speed scores may be a factor in improved WCST performance.) Taken en masse, these data suggest that the students, ages 8-11 in this referred sample demonstrated significantly more difficulties in the self-regulation executive control of reasoning, attentional, and shifting capacities than those in the other two age groupings, as evidenced by their performance on the WCST. Given the high number of students with the diagnosis of ADHD or the likely diagnosis of ADHD, this result appears to be similar to the findings of Barkley, Grodinsky, and DuPaul (1992), which suggested that WCST children's performance improved with age in the ADHD proband.

Question 3: Does a process-oriented approach to WCST performance and patterns of responding provide important clinical information that is not captured in the WCST score indices?

Given the complex nature of the WCST, it would seem that an adequate analysis of performance requires the use of a process approach to interpretation, whereby not only the patterns of responses but also the observations of the student during WCST performance are analyzed. The same process approach principles that were delineated in the development of the WISC-IV Integrated (McCloskey & Maerlander, 2005) could easily be applied in WCST interpretation:

1. Students can obtain similar scores on WCST variables, yet demonstrate different patterns of responding.
2. Individual scores are a likely reflection of the integration of multiple capacities including various executive function components (for example, although Perseverative Errors is thought to be a measure of

shifting capacities, other executive functions are likely contributing to performance on this WCST variable).

3. In observing a student's performance, the clinician can observe specific strategies and behaviors that the student demonstrates (for example, the degree to which a student demonstrates emotional self-regulation or lack thereof is likely to impact performance).

Specific lines of inquiry for exploring this research question include: Does the pattern of responses offer more data relative to variables that contribute to performance and impairment, as suggested by Heaton et al. (1993)? Can clinically relevant information be obtained from the number of trials to complete all WCST sets, not only the first set? Are there clinically relevant variations in the response patterns in cases in which comparable WCST variable scores are obtained? When are observations of WCST response patterns helpful in understanding difficulties that are not revealed in the WCST score indices?

As suggested by the data relative to number of trials completed (m = 5.22), most students in this sample were able to complete the full requirements of the tasks, i.e. match the cards by the three dimensions over two rotations. However, there was considerable variability in the number of trials needed to complete sets, especially in the first three sets. Table 8 shows the mean scores for set completion by trials. The greatest variability in performance is evident in the first two trials, in which the mean number of trials was 14.43 and 24.18 and the range for completion extended from 10 trials (meaning all of the first ten attempts matched the prevailing pattern of color or form) to 86 and 108. For the first set, however, only 4 of the 91 cases for which the category breakdown was available required more than 25 trials to complete the first set. As suggested by the mean of 24.18 for the second set, the shift to a different set (the concept of form) was more difficult in many cases. Although there were 35 students who were able to complete the second set in 15 trials or less, there were 25 who required 25 or more trials to complete that set, as well as one student who was unable to complete the set. It appears that the concept of set-switching was established for most after this point, because the mean number of trials to complete the third set dropped to 17.28, with a range of 10 to 62 trials. Data for the switch back to the initial matching principle produced a mean and range comparable to those obtained for the third set. Not surprisingly, for those able to complete the test, the fifth and sixth sets produced the lowest means and ranges.

A cross-tabulation of number of trials to complete categories by other scoring variables again suggested little relationship between variables. Trials to complete categories were recoded into category variables: 1 = 10-14 trials; 2 = 15-19; 3 = 20-24; 4 = 25-29; 5 = 30-39 and 6 = > 39 trials. Category coding for standard scores was the same as noted in the previous sections. For those rating scale t-scores, variables were recoded as: 1 = t-scores < 30; 2 = 30-39; 3 = 40-49; 4 = 50-59; 5 = 60-64; 6 = 65-70; 7 = 70-79; 8 = 80 or greater. All of the students who scored less than 90 on Perceptual Reasoning (n = 4), Verbal Comprehension (n =4) and Matrix Reasoning (n = 7)

Table 8

Descriptive Statistics for Trials to Complete Categories

Category Sets	*n*	*M*	*SD*	Min	Max
First	91	14.43	11.79	10	86
Second	90	24.18	20.13	10	108
Third	83	17.28	11.61	10	62
Fourth	76	17.07	9.06	11	62
Fifth	72	14.15	5.73	11	37
Sixth	65	14.26	4.12	10	27

completed the first category in fewer than 15 trials. However, of those who scored higher than 109 on Perceptual Reasoning, 7 of 38 (18%); on Verbal Comprehension, 8 of 49 (16%); and on Matrix Reasoning, 7 of 48 (15%), required more than 14 trials to complete the first category. Three of the 14 (21%) students who scored less than 90 on Working Memory required more than 14 trials to complete the first category, with 3 (15%) requiring 20 or more trials, but 16 of the 18 (89%) students who scored above 109 on Working Memory completed the set in 14 trials or less. Those with low or high Processing Speed scores tended to do well on trials to complete the first category, with only 3 of the 24 (12%) scoring below 90, and 2 of the 11 (18%) scoring above 109 requiring more than 14 trials.

Table 9 shows the distribution of number of trials to complete the second category by Wechsler variables. The results suggest that those who fell below 90 on Perceptual and Matrix Reasoning were able to complete the second set within 19 trials, but the number of trials needed for those scoring equal to or above 90 on Perceptual Reasoning tended to vary considerably. Also, almost all of the students tended to complete the 2nd category in fewer than 20 trials; however, those with higher (greater than 109) Verbal Comprehension Index scores varied widely in the number of trials required. Cross-tabulations based on Working Memory and Processing Speed score ranges reflected greater variation in the relationship between numbers of trials for completion and Wechsler score category. Although the number of students in the lower cognitive ranges was few, Table 9 certainly suggests that having weaker cognitive capacities is not particularly related to the ability to establish and maintain the second category of the WCST.

Cross-tabulation of behavior rating scale variables by number of trials to complete categories yielded similar distributions of scores. Because the shift to the second matching pattern appears to require more of the executive functioning demands of set-shifting, inhibition, and working memory, the distribution of scores in relation to trials to complete the second set may be particularly interesting. Table 10 provides data relative to BRIEF Shift, Inhibit and Working Memory scales. Again, the only variable that appeared to indicate greater dispersion of scores was the Working Memory responses of parents, because one-third (11 of 33) of the students who required more than 19 trials for completion of the second set had parent ratings of 65-79.

Table 9

Distribution of Wechsler Scores by Number of Trials to Complete Second WCST Set

	Number of Trials					
Wechsler Variable	10-14	15-19	20-24	25-29	30-39	> 39
Perceptual Reasoning						
Less than 90 (*n* = 4)	2	1	0	0	0	1
90-99 (*n* = 16)	2	6	2	1	2	3
100-109 (*n* = 21)	8	5	1	3	2	2
Greater than 110 (*n* = 37)	19	8	2	3	3	2
Verbal Comprehension						
Less than 90 (*n* = 4)	1	1	0	1	0	1
90-99 (*n* = 7)	4	1	0	1	0	1
100-109 (*n* = 19)	5	5	2	0	3	4
Greater than 110 (*n* = 48)	21	13	3	5	4	2
Matrix Reasoning						
Less than 90 (*n* = 7)	7	0	0	0	0	0
90-99 (*n* = 6)	5	1	0	0	0	0
100-109 (*n* = 20)	16	2	1	1	0	0
Greater than 110 (*n* = 48)	41	2	3	0	1	1
100-109 (*n* = 17)	9	5	2	0	1	0
Greater than 110 (*n* = 18)	8	5	1	1	1	2
Working Memory						
Less than 90 (*n* = 14)	6	2	1	1	2	2
90-99 (*n* = 28)	7	9	1	4	3	4
Processing Speed						
Less than 90 (*n* = 24)	5	7	2	3	3	4
90-99 (*n* = 23)	12	7	0	2	1	1
100-109 (*n* = 20)	8	3	3	1	3	2
Greater than 110 (*n* = 11)	5	4	0	1	0	1

Table 10

Distribution of BRIEF Parent and Teacher Scores by Number of Trials for Second Set

	Number of Trials					
BRIEF Variable	10-14	15-19	20-24	25-29	30-39	> 39
Parent Inhibit						
Less than 49 (*n* = 22)	10	6	3	1	1	1
50-64 (*n* = 30)	11	12	1	5	1	0
65-79 (*n* = 14)	4	3	1	0	2	4
Greater than 79 (*n* = 6)	3	1	0	0	1	1
Teacher Inhibit						
Less than 49 (*n* = 12)	4	5	0	3	0	0
50-64 (*n* = 22)	5	7	3	1	3	3
65-79 (*n* = 7)	1	4	0	0	1	1
Greater than 79 (*n* = 4)	2	1	0	0	0	1
Parent Shift						
Less than 49 (*n* = 18)	7	9	1	1	0	0
50-64 (*n* = 26)	13	3	4	3	0	3
65-79 (*n* = 22)	6	8	0	2	4	2
Greater than 79 (*n* = 6)	2	2	0	0	1	1
Teacher Shift						
Less than 49 (*n* = 7)	1	4	1	0	0	1
50-64 (*n* = 19)	8	6	1	2	1	1
65-79 (*n* = 12)	1	5	1	1	2	2
Greater than 79 (*n* = 7)	2	2	0	1	1	1
Parent Working Memory						
Less than 49 (*n* = 8)	6	2	0	0	0	0
50-64 (*n* = 25)	8	9	3	2	0	3
65-79 (*n* = 33)	11	9	2	4	5	2
Greater than 79 (*n* = 6)	3	2	0	0	0	1
Teacher Working Memory						
Less than 49 (*n* = 2)	0	2	0	0	0	0
50-64 (*n* = 14)	4	5	1	2	1	1
65-79 (*n* = 16)	5	5	1	1	2	2
Greater than 79 (*n* = 13)	3	5	1	1	1	2

As noted by Heaton et al. (1993), the pattern of performance on the WCST may yield relevant information that is not evident in the scoring indices. This is especially true when one looks at students who received comparable WCST scores, but who obtained these scores through very different patterns of performance. As previously discussed, a process approach which examines the patterns of responses and the behaviors exhibited may be diagnostically relevant, particularly as it relates to the multiple executive capacities that are or are not utilized in WCST performance. This point was exemplified through the review of students with almost identical Error Percentage and Conceptual Level Response scores on the WCST.

Table 11 provides the scores for four children who performed poorly (i.e., scoring below 80), three children who performed in the average range, and three children who performed in the superior range on WCST Error Percentage and WCST Conceptual Level Responses. None of the students who performed in the below average group, not surprisingly, was able to meet the full requirements of the task because none completed 6 categories. However, although their Error Percentage and Conceptual Level scores were virtually identical, two of the four completed three categories; one, 2 categories; and 1, only 1 category. Student A, a thirteen year old girl, was able to discern the correct matching pattern with her first attempt and thus completed the first category within ten trials. However, she had significant difficulty making the shift to the second category, requiring 34 trials. She did appear to discern the second matching dimension by matching for form in three and four successions, but was not able to maintain set during the first 24 of the 34 trials, suggesting lapses in attention. When she made errors in this area, they tended to be quite perseverative, often continuing to match for the first dimension, suggesting the kind of –return to set" perseveration that was discussed in the

Table 11

Patterns of Performance of Students with Range of WCST Variables

Student	Err%	Per Err%	NonPer Err%	ConL Resp%	CatCom	1st	2nd	3rd	4th	5th	6th
Below Average:											
A	77	91	73	75	3	10	34	18			
B	78	91	72	75	2	11	24				
C	73	59	98	75	3	13	45	10			
D	72	73	76	73	1	86					
Average:											
E	100	107	91	101	6	12	19	12	33	18	12
F	103	110	95	102	3	11	27	53			
G	102	107	98	104	6	11	17	20	28	27	12
Above Average:											
H	136	>145	110	>145	6	19	22	11	11	11	11
I	132	>145	117	130	6	11	10	14	13	11	11
J	129	145	NA	123	6	12	12	10	11	16	12

Note. WCST Err% = Error Response Percentage; WCST PerErr% = Perseverative Error Percentage; WCST NonPerErr% = NonPerseverative Error Percentage; WCST ConLRes% = Conceptual Level Response Percentage; WCST CatCom = Categories Completed

literature review. Although she was able to discern the third matching pattern—number of design elements on each card—in 18 trials, she was never able to make the shift back to the original dimension of color, even though she had three or four correct responses on three separate occasions. Because Failure to Maintain Set is determined after five or more correct responses, none of these set “losses” were characterized as such, and the only supplemental score that was below average was Learning to Learn (2-5%). Student A‗s cognitive profile suggested at least average verbal comprehension and perceptual reasoning capacities; on the WISC-IV, her Matrix Reasoning score was a 10. Although not identified as having ADHD, her parents and teachers noted concerns relative to sustained attention. BRIEF parent ratings were significant for difficulties with cognitive flexibility and emotional control. As Student A was completing assessment tasks, she noted, ―You‗re making me think. I don‗t like thinking!" suggesting that she struggled with sustaining attention and/or effort.

Student B, a nine year old girl, was easily able to complete the first set in 11 trials and required 24 trials to complete the second set, but was never able to discern and match the cards to complete another set. The supplemental scoring variables were somewhat indicative of her difficulties here as she fell in the 6-10 percentile range for categories completed and the 2-5th percentile for Learning to Learn. In discussing her behaviors during the testing, the psychologist noted that ―When applying reasoning to more abstract nonverbal visual content, [this student demonstrated] a lack of consistency of her efforts and the tendency to rush through working out solutions without monitoring for accuracy." Student B‗s cognitive functioning fell in the above average range, with especially high Perceptual Reasoning (119). Processing Speed was well below average (SS = 67) and Working Memory, low average (SS = 87). Her classroom teacher noted concerns regarding attention problems, the ability to shift cognitive set, task initiation and persistence, and task planning and monitoring.

Student C, a 14-year-old male, also had relatively little difficulty completing the first pattern in 13 attempts, but required 35 trials to determine the second matching pattern. In completing the Form pattern which included the two final cards which also matched for Number, he appeared to shift to this new dimension and thus unknowingly transitioned successfully to complete the third matching pattern. However, when it was time to switch back to the initial pattern of Color, he was unable to do so in the remaining 60 trials. As noted by the evaluator, ―He was unable to sustain his attention and effort long enough" to complete the task. A Learning to Learn score could not be computed (because he had not completed a sufficient number of categories) and Student C's Categories Completed score fell at the 11-16%. Although Student C indicated borderline overall cognitive and verbal abilities, his Perceptual Reasoning score was a 90; Matrix Reasoning was a 7. Behaviors of concern included significant difficulties in maintaining attention and sustaining effort to tasks of more than 20 seconds in duration, in using working memory skills, in attending to details, and in retrieving previously learned information.

Student D, a nineteen year old male, required 86 trials to complete the first and only set completed. He, as with student C, was able to match correctly in 5-8 successions; however, he was not able to sustain the matching principle through to the requirement of 10 successive matches until the 86th attempt. He was never able to discern the second matching pattern, often returning instead to the first matching dimension, again demonstrating perseverative responding. Frequently he would match a card correctly, only to make an error on the next attempt. The totality of his scores, included supplemental scores, are reflective of his degree of difficulty, but not the degree to which he appeared to struggle with the multiple factors required in WCST completion. This student's overall cognitive functioning fell into the above average range (FS = 116), with superior verbal comprehension and a Matrix Reasoning score of 13. In spite of this strong cognitive profile, his grades in high school ranged from A to D and he now was in jeopardy of being asked to withdraw from his college program because of poor grades. In discussing the difficulties with this test, the student explained that he was attempting to discern patterns which involved dual-dimension matches or matches which involved the sequence in which the cards appeared, never discerning that the simple match of color, shape or number was indicated. In reviewing these four cases the difficulties are hinted at with the supplemental scoring variables, but the different patterns of their difficulties do not appear to be well differentiated by the WCST scores themselves.

The three students who demonstrated average performance on the WCST scoring variables also displayed different patterns of performance. Although two of the three who received average scores were able to complete the full requirement of the task because they completed 6 categories, one was able to complete only three categories. All three were able to establish and complete the first set with relatively few (11-12) trials, although all required considerably more attempts to complete the second, with one requiring as many as 27 trials. Student E, a ten year old boy, completed the first set in 12 trials, but required 19 trials to complete the second. In his case, in spite of being provided feedback that his responses were incorrect, he attempted to use the initial match concept for 6 of 9 trials. He was then able to sustain to complete the second set and required only two attempts before discerning the matching dimension for the third set. He had difficulty switching back to the first matching pattern, perseverating on the number dimension. He had less difficulty completing the fifth set (18 trials) and easily completed the sixth (11 trials). The supplemental scoring indicated no difficulties, but this student in fact demonstrated a perseverative response pattern that significantly affected performance. Student E's Wechsler scores were solidly in the average range, with a Matrix Reasoning score of 10. Parent and teacher BRIEF responses were suggestive of difficulties with planning, organizing, and sustaining; parent and teacher responses as well as observations during testing suggested difficulties with shifting set.

Student F, also a ten year old boy, completed the first set in 11 trials. He then demonstrated lapses in attention while attempting to complete the second category, losing set twice after providing multiple correct responses and requiring 27 responses to complete the second set. In switching to the third dimension, he needed 25 responses before seeming to discern the third pattern and lost set twice, each time after 8 successive correct matches. Ultimately he required 53 responses to establish the third set and never completed the fourth set, breaking set after 9 correct responses and then again after 6 correct responses. Assessment of Student F's cognitive abilities suggested very superior (142) verbal capacities and above average visual/spatial and visual reasoning abilities, with a Matrix Reasoning score of 15. Processing Speed and Working Memory indices were more average. BRIEF and BASC parent ratings were significant for attention problems and difficulties with planning and organization. In testing Student F, the psychologist noted that when he was given more complex and challenging tasks, Student

F demonstrated strong engagement and focus, but when engaged in simple and routine tasks, he had difficulty in maintaining attention and effort.

Student G, another ten year old boy, had no difficulty establishing the first set, requiring only 11 trials to do so. When required to shift set, his performance dropped somewhat; he required 7 trials to make the shift and 17 trials to completed the second set. He performed similarly with the third set, needing 10 trials to make the shift and 20 trials to complete the third set. He quickly discerned the shift back to the original matching pattern; however, it was here that he began to demonstrate lapses in attention, breaking set twice en route to completing the fourth set in 28 trials. The shift to the fifth set required 17 trials, but the shift to the sixth set required only two trials. Student G was initially referred to the psychologist conducting the evaluation to determine if he might be eligible for the district's gifted and talented program. Although Working Memory and Verbal Comprehension skills were in the superior to very superior range, his Perceptual Reasoning score was a 108; Matrix Reasoning was a 10. Teacher responses to the BASC-2 noted concerns relative to attention problems and anxiety; BRIEF responses were significant for difficulties with inhibition, planning and organizing, and organization of materials. Although two of the students performing in the average range demonstrated difficulties in sustaining attention to task completion, only Student F had WCST scores (Learning to Learn, 11-16 percentile; Failure to Maintain Set below the first percentile and Categories Completed, 2-5 percentile), that were indicative of those difficulties.

Although the selected cases that performed well on the WCST displayed similar performances because all categories were completed and most of these categories were completed within a relatively few number of trials; variations in performance also were evident. Student H demonstrated relative difficulty in establishing the first set (9 trials to establish, 19 to complete) and shift to the second set (12 trials to shift, 22 trials to complete), but then made very quick transitions for the remaining four sets, requiring only 1 trial for each shift. The performance pattern for this fourteen year old female may be suggestive of some relative difficulties with task initiation. Referred because of her difficulties in initiating and completing academic tasks, Student H‗s cognitive profile suggested strong reasoning capacities (Verbal Comprehension score of 121, and Perceptual Reasoning, 115 and a Matrix Reasoning score of 16), but more average Working Memory (102) and Processing Speed (100). BRIEF teacher responses suggested difficulties with task persistence and sustaining (under Working Memory) and monitoring performance; parent responses suggested concerns regarding all scales of the BRIEF as well as endorsed depression, hyperactivity and withdrawal as clinically significant on the BASC-2.

Student I quickly discerned the first pattern (1 trial to establish, 11 to complete), made a seamless shift to the second set (0 trials to shift, 10 trials to complete), required additional attempts (3 trials and 4 trials) to complete sets three and four and required only 1 trial each to complete the shifts for the remaining two sets. Student I‗s cognitive evaluation suggested very superior (132) Verbal Comprehension and superior (127) Perceptual Reasoning scores, with a Matrix Reasoning score of 16. Processing Speed fell in the average range (109) and Working Memory, high average (113). In spite of his strong cognitive functioning, his written expression scores were below average and he indicated difficulties with verbal fluency and ability to shift cognitive set.

Student J was very efficient with all but the fifth set, appearing to have a relative lapse in attention, requiring 6 trials to make the shift. What distinguished this student's performance from that of others who did well was the extreme speed at which she completed the task; the entire WCST performance required only about 5 minutes. Student J also indicated superior cognitive functioning, with a Perceptual Reasoning score of 131 and a Matrix Reasoning score of 15. Although she had been prescribed medication for diagnosed ADHD symptomotology, Student J nonetheless demonstrated difficulties in sustained focus and task persistence. At times, the speed at which she completed tasks resulted in a positive outcome, such as on the WCST; however, her overall school performance was hampered by the ADHD deficits.

Although not clearly evident from the formal score indices of the WCST, a process approach analysis of the response patterns of these 10 cases suggests that additional information of clinical value could be gleaned from their performances. A review of the response patterns, particularly the difficulties in making the shifts from one matching dimension to another <u>after</u> the first set is completed; the loss of set after three or more consecutive correct matches; the perseverative response patterns involving reverting to a prior set or a ―stuck-in-set" response pattern in spite of negative feedback, appear to be indicators of some kinds of specific executive function weaknesses that have been cited in the literature. The results of this kind of process approach analysis suggest that examining response patterns can offer a valuable perspective on executive functioning difficulties that are not necessarily revealed by the WCST score indices, including the supplemental scores.

Chapter Four

Discussion

Summary of Results

The Wisconsin Card Sorting Test (WCST) is widely used for neuropsychological assessment of executive functions. Although the literature notes that the WCST is a measure of abstract reasoning and cognitive flexibility, there has been little data relative to those constructs that are assessed when the test is used with children or to the relationship between WCST performance and performance on other child assessment tools. The purpose of this study was three-fold. First, it was designed to investigate the relationship between student performance on the WCST with other clinical measures developed to assess similar executive function or cognitive capacities. Second, the study investigated factors or clusters of factors, such as age, gender, diagnosis, or test administration format, to determine if any of these differentially impacted WCST results in children and adolescents. Third, this study looked at patterns or observations of WCST performance to determine whether or not such a process approach could provide clinically relevant information about student functioning that is not revealed in the scores that the student may obtain.

As hypothesized, the findings relative to the first question suggested very low correlations between scores and performance indicators from the WCST and measures of cognitive capacities, behavior ratings of parents and teachers, and executive function measures that involve task switching, response inhibition, and problem-solving. In fact, the results of correlations and most cross-tabulations found a wide distribution of scores: the number of students who performed well on Wechsler variables and did well on WCST variables was often equally matched by the number of students that did well on the Wechsler but poorly on the WCST or poorly on the Wechsler scales and well on the WCST. Although more students did better on the WCST than on D-KEFS Color-Word Interference Inhibition and Inhibition/Shifting, the differences were not significant and scores on D-KEFS Twenty Questions Asked were widely distributed among WCST scores. Rating scales appeared to be a particularly poor predictor of WCST performance, reflecting the weakest correlations of all the assessment tools in the study.

A closer analysis of the data did, however, suggest an interesting relationship between measures of reasoning on the WCST and Wechsler Matrix Reasoning. Seventy-five percent of all the students in this data set and 92% of those who had scores of 12 or higher on Matrix Reasoning had scores on WCST Conceptual Level Responses that were equal to or lower than their Matrix Reasoning scores. Thus, as noted previously for this data set, it was difficult to perform better on Conceptual Level Responses than on Matrix Reasoning. The relationship between Matrix Reasoning and Conceptual Level Response scores suggested here may be reflective of the association frequently noted in the literature review between cognitive abilities and executive capacities.

When WCST performance was investigated relative to subgroups, the results yielded no gender or diagnosis differences. There were, however, differences in performance that was based on how the test was administered; those who were administered the computer version tended to earn higher non-perseverative response scores, but those who were administered the manual version earned higher perseverative error scores. Further analysis suggested that the Perseverative Error difference might have been due to a sampling artifact of this data set, so these findings are to be viewed cautiously. Distributions were less skewed for Non-Perseverative Errors, suggesting that for this population, the form of WCST administration influenced engagement in random responding, with those using the manual version having poorer scores. Also, of interest, was the finding indicating that in spite of age-corrected scoring, the oldest group of students in this sample out-performed the youngest on all the scoring variables, and an intermediate age group outperformed the youngest in Perseverative Errors and Categories Completed. Although these results may be counter-intuitive, given the scoring mechanism, it is confirmatory of other literature (Barkley et al., 1992) that suggested that WCST performance of ADHD children (who are significantly represented in the sample) did improve with age. When age-group correlations of WCST variables with other clinical measures were conducted, results suggested a somewhat stronger relationship of WCST performance with processing speed in younger children, although other analyses in this study suggested that speed of processing had a different impact in older students. That processing speed may be a factor in WCST performance is also suggested by previous findings (Salthouse, 2005).

The third question of this study examined patterns and observations of WCST performance that appeared to offer information about executive function capacities that were not readily revealed in the formal scoring mechanisms. Although the responsible investigator noted previously the observation that students appeared to have more difficulty making the shift back to the original matching pattern (the fourth category), the results from this study suggested instead that the shift to the second category was the most difficult. Although formal WCST scoring processes provide normative data for the number of trials to complete the first category set, the results of this study suggest that the number of trials to complete the second category may be the more interesting and clinically relevant index, because the range in the number of attempts to complete the second set varied considerably in this referred population, with the mean number of trials being significantly higher for the second set than the mean number of trials for the first set and for all of the other remaining sets.

Significance of the Results

Although there are considerable studies that have reported strong correlations between WCST variables and those on measures of cognitive abilities, such as the Wechsler scales, there also have been studies reporting minimal correlations. This study falls into that latter category indicating that cognitive measures, behavior rating scales and other executive function measures, such as those found on the D-KEFS, are not likely to predict WCST performance in children. Because the correlations among all of the various measures in this study were low, the results suggest a dissociable nature for each of these assessment tools, an idea which parallels some of the more recent models of executive functions (e. g., McCloskey et al., 2009). Although WCST, Wechsler and D-KEF‘s tasks involve the use and manipulation of symbolic material, each appears to measure unique combinations of executive and cognitive capacities.

Of particular note, given the fact that the WCST was designed as a test of abstract reasoning, was the relationship between measures of abstract reasoning on the Wechsler and WCST Conceptual Level Response scores, also presumed to be a measure of abstract reasoning. As noted in the literature review, many discussions seem to suggest that reasoning and executive functions are similar or overlapping cognitive constructs. Overall, the data from this study indicate that children and adolescents scoring below average on Wechsler Verbal Comprehension and Perceptual Reasoning tasks are as likely as those who have above average cognitive scores to score well on the WCST indices.

However, analysis of the relationship between Wechsler Matrix Reasoning and WCST Conceptual Level Response scores suggests that they may represent a continuum of reasoning capacities. For the most part, students in this data set were not able to out-perform their basic reasoning capacity, as measured by Matrix Reasoning, in the completion of the WCST, as measured by their Conceptual Level Response score. Results suggest that the WCST may be a measure of the executive function **direction** or **control** of reasoning capacities rather than an assessment of the reasoning capacities themselves. A review of the task requirements of each may be helpful in understanding this relationship. Although some self-regulation executive function capacities are needed for Matrix Reasoning performance because the student has to initiate the task and hold the sequence and alternatives in mind while selecting one of the five options given to complete the matrix, the executive function demands are minimized through the provision of explicit directions, defined options, and practice about how to perceive, think about, and respond to the task. On the WCST, the student is not provided with the explicit directions or correct option. The student is told that he or she needs to match the cards, but how the match is made, how many times the match needs to be made, and what shifts are required is determined by the feedback that the student is given. He or she needs to self-regulate his or her own reasoning capacities to make these determinations. Certainly, observations of WCST performance suggested that at times students in this data set who have very strong cognitive abilities, such as Student D, found it difficult to discern the matching principle by looking for a much more complicated solution for card sorting. Although he understood the reasoning requirement for the task, he had significant difficulty in determining the demands of the WCST, given the ambiguous information provided, demonstrating difficulties with the regulation of the executive capacities that were required to solve the problem. The contrast in student performance between these two measures appears to indicate a change in effectiveness based on task demands, with WCST results for most students in this data set representing decremental production because of the increased need for self-regulation cueing. Each measure then provides valuable, but unique information. Use of both measures, Matrix Reasoning and WCST Conceptual Level Responses, may provide what Delis et al. (2001) indicated was critical, i. e., “the empirical measures for determining whether poor performance is due to deficits in more fundamental cognitive skills or deficits in…executive functions” (p. 3).

Both the WCST and the D-KEFS tasks involve problem-solving, inhibition of response and set-shifting with symbolic material; however, the variation in task demands as well as the differing information processing demands appears to contribute to variability in level of performance. Again, in completing the D-KEFS executive functioning tasks, the student is provided the opportunity to practice the task, having been given directions that are highly specific, offering clear delineation about how the task should to be completed. Although some examinees benefit from the instructional experience and might not perform well either on the D-KEFS or on the WCST without such instruction, others might benefit from the lack of specific instruction to the WCST, perceiving the ambiguous nature of the task as representing more of a challenge or puzzle, and as a result, more interesting or motivating.

Behavior rating scales offer judgments about a child or adolescent‘s use of executive function capacities similar to those assessed by the WCST and the D-KEFS, but these scales assess behavior across the multiple arenas of use of symbolic materials, interpersonal and intrapersonal functioning, and interaction with the environment. Despite the apparent overlap in terms of the executive function constructs assessed, behavior ratings of executive functions directing multiple domains of functioning (perception, emotion, cognition and action) utilized across multiple arenas of involvement showed little relationship with the examinees‘ performances with the WCST, suggesting that one is not likely to be able to predict WCST performance from rating scale results. These results also are consistent with some of the conceptual models of executive functioning capacities (e. g., McCloskey et al., 2009) that suggest that executive functions may be arena specific; i.e., the level of effectiveness of executive function direction of work with symbolic materials is not necessarily related to the level of effectiveness of executive function direction in the interpersonal, intrapersonal or environmental arenas.

Results from this study also suggest that the capacity to discern, maintain, and complete the second category of the WCST may reflect the core construct of cognitive flexibility. It is in the shift to the second set that the student is likely to become aware of the new dimension to this assessment tool; access to problem-solving strategies is required in order to discern the shift to a new response pattern. One-quarter of the students in this study found the shift to the second set to be highly difficult and required 25 or more trials to complete the set. This number represented an amount that was significantly higher than that required for completion of any of the other sets. Observations of students in this study who had difficulty with the shift to the second set also often revealed a loss of set rather than simply an inability to discern the shift. However, because loss of set was apparent in all other category trials, the increased number of trials to complete the second set may be in fact a significant reflection of initial set-shifting difficulties, and as such, an important scoring dimension not currently reflected in the Heaton system.

This study showed that the analysis of the patterns of performance and observations related to performance of individual students can provide clinically relevant qualitative information about executive function capacities that is not adequately reflected in the existing WCST scoring dimensions. Beyond the pattern of number of trials to complete sets discussed above, patterns of performance and specific observations may effectively identify a loss of set in fewer trials than the Heaton definition of 5 or more successive correct responses prior to an error. It is interesting to note that understanding of set, as defined by the Heaton scoring for Conceptual Level Responses, is three successive correct matches. Process-oriented analysis of performance of the students in this study suggested that loss of set may occur at the same point of set understanding, not two matches later.

Limitations of the Study

This study has several limitations. This is a nonexperimental research design, which limits the generalizability of the findings. The sample size was relatively small and limited to a certain geographic and demographic sampling because most of the participants came from Fairfield County, Connecticut, towns, from several suburban counties in New York, as well as from a defined region in southeastern Pennsylvania. Although it was a “referred” population, the reasons for referral, educational classification, and diagnoses or lack thereof were diverse and the test variables revealed a particularly wide scoring range which at times was reflected in poor homogeneity of variance among the entire group; this again is likely to limit the degree to which the findings may be generalizable. Additionally, medications effects were not monitored, which, as suggested by Hale et al. (2005), may impact degree of measured neuropsychological impairment.

Because the students in this data set were administered assessment instruments as part of a school-mandated or private psychological evaluation, there was no control over which tests or subtests would be included in the assessment battery. As a result, comparisons of assessments were dependent on the data provided rather than on comparisons that might have contributed systematically to a research study. Also, although the study used data from as far back as 2000, data from cases using the WISC-III was not included in the analyses; use of these would likely impact generalizability.

Although administration and scoring procedures for all the instruments used in the study are standardized, manual administration of the WCST, in particular, requires significant training in scoring procedures and practices; thus, with the collection of data from multiple school psychologists, there may be some scoring variability. Also, as is also suggested by reliability data from the normative sample, there may be low interrater reliability on data from the rating scales completed by parents and teachers; this is likely to contribute to the results relative to these behavior rating scales.

Contributions to the Field of School Psychology

With the increasing use of neuropsychological assessment tools by school psychologists and the emerging understanding of how executive function deficits impact student performance, it is more important for school psychologists to be able to utilize assessment tools that provide relevant information about students‘ functioning. Because the role of the school psychologist is also to provide recommendations for effective interventions to improve student functioning, it is important that the assessment data are able to differentiate between cognitive and executive function deficits and to delineate the arena in which these deficits are likely to occur. This study suggests that neuropsychological assessments and behavior rating scales purport to measure similar constructs; however, use of one is not likely to provide all the information needed to have a good understanding of an individual‘s executive function capacities. It then becomes incumbent upon psychologists to understand the contributions as well as limitations of each instrument used. Of the executive function measures involving the use and manipulation of symbolic material, however, the WCST in particular, with its ambiguous directions and multifactorial nature, may offer specific and unique information relative to an individual's self-regulation execution function of reasoning. The results also support the use of the WCST as a measure of the capacity to cue and direct cognitive flexibility (set-shifting) and sustained attention, especially when the examiner employs a process approach to the interpretation of test performance and results. Thus, the WCST may be a particularly useful tool in an executive functioning assessment battery when difficulties related to symbol system use are reported.

Future Directions

Given not only the questions emerging from this study, but also the limitations of the study, recommendations for future research include the following:

1. A larger sample size with a uniform assessment battery to explore in a greater in-depth manner, the relationship between reasoning scores on the Wechsler and scores on the WCST.
2. Although the results of this study suggested that specific executive capacities improved within this age range (8-19), these findings warrant further study to understand more fully which specific executive function capacities improved to allow for the stronger WCST performance. Further study on the impact of speed of processing on WCST performance by age group is also suggested.

3. Given the fact that this study suggests that the number of trials to complete the second WCST set may be more reflective of set-shifting capacities, development of normative data for this variable as well as further study relative to number of trials to complete the second set and its relationship to other measures of cognitive flexibility may be warranted.

4. Further study relative to whether or not administration format does differentiate WCST performance is suggested. Because the numbers of this data set are fairly small, the results relative to administration format may artificially suggest weaker inhibition of response or lapse of attention in children taking the manual version of the WCST.

References

Alvarez, J. A. & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*(1), 17-42.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. Text Revision) Washington, D. C.: Author.

Arffa, S., Loveli, M., Podell, K., & Goldberg, E. (1998). Wisconsin Card Sorting Test performance in above average and superior school children. *Archives of Clinical Neuropsychology, 13*(8), 713-720.

Avila, C., Cuenca, I., Felix, V., & Parcet, M. (2004). Measuring impulsivity in school-aged boys and examining its relationship with ADHD and ODD ratings. *Journal of Abnormal Child Psychology, 32*(3), 295-305.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General, 130*(4), 641-657.

Baddeley, A., Della Sala, S., Papagno, C., & Spinnler, H. (1997). Dual-task performance in dysexecutive and nondysexecutive patients with frontal lesion. *Neuropsychology, 11*(2), 187-194.

Baldo, J. V., Shimamura, A. P., Delis, D. C., Kramer, J., & Kaplan, E. (2001). Verbal and design fluency in patients with frontal lobe lesions. *Journal of the International Neuropsychological Society, 7*, 586-596.

Barcelo, F. (2000). Does the Wisconsin Card Sorting Test measure prefrontal function? *The Spanish Journal of Psychology, 4*, 79-100.

Barkley, R. A. (1996). Linkage between attention and executive functions. In G. R. Lyon & N. A. Krasnegor (Eds.), *Atttention, memory and executive function* (pp. 301-325). Baltimore: Paul H. Brookes.

Barkley, R. A. (1997). *ADHD and the nature of self-control.* New York: Guilford Press.

Barkley, R. A., Grodzinsky, G., & DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: A review and research report. *Journal of Abnormal Child Psychology, 20*(2), 163-188.

Baron, I. S. (2004). Test review: Delis-Kaplan Executive Function System. *Child Neuropsychology, 10*(2), 147-152.

Baxter, L. R., Clark, E. C., Iqbal, M., & Ackermann, R. (2001). Cortical-subcortical systems in the mediation of Obsessive-Compulsive Disorder: modeling the brain‘s mediation of the classic ―neurosis.” In D. G. Lichter & J. L. Cummings (Eds.), *Frontal-subcortical circuits in psychiatric and neurological disorders* (pp. 207-230). New York: Guilford Press.

Borkowska, A., & Rybakowski, J. K. (2001). Neuropsychological frontal lobe tests indicate that bipolar depressed patients are more impaired than unipolar. *Bipolar Disorders, 3*(2), 88-94.

Borkowski, J. G., & Burke, J. E. (1996). Theories, models and measurements of executive functioning: an information processing perspective. In G. R. Lyon and

N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 263-278). Baltimore: Paul H. Brookes.

Carrillo-de-la-Pena, M. T. & Garcia-Larrea, L. (2007). Right frontal event related EEG coherence (ERCoh) differentiates good from bad performers of the Wisconsin Card Sorting Test (WSCT). *Clinical Neuropsychology, 37*, 63-75.

Chelune, G. J. & Thompson, L. L. (1987). Evaluation of the general sensitivity of the Wisconsin Card Sorting Test among younger and older children. *Developmental Neuropsychology, 3*(1), 81-89.

Decker, S. L., Hill, S. K., & Dean, R. S. (2007). Evidence of construct similarity in executive functions and fluid reasoning abilities. *International Journal of Neuroscience, 117,* 735-748.

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

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

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

Denckla, M. B. (1996). A theory and model of executive functions: A neuropsychological perspective. In G. R. Lyon and N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 263-278). Baltimore: Paul H. Brookes.

Denckla, M. B. (2007). Executive function: binding together the definitions of Attention Deficit/Hyperactivity Disorder and learning disabilities. In L. Meltzer (Ed.), *Executive function in education: From theory to practice* (pp. 5-18). New York: Guilford Press.

Fischer, K. W., & Daley, S. G. (2007). Connecting cognitive science and neuroscience to education: potentials and pitfalls in inferring executive processes. In L. Meltzer (Ed.), *Executive function in education: From theory to practice* (pp. 5-18). New York: Guilford Press.

Fisher, N. J., DeLuca, J. W., & Rourke, B. P. (1997). Wisconsin Card Sorting Test and Halstead Category Test performances of children and adolescents who exhibit the syndrome of nonverbal learning disabilities. *Child Neuropsychology, 3*(1), 61-70.

Gioia, G. A., Espy, K. A., & Isquith, P. K. (1996). *Behavior Rating Inventory of Executive Function—Preschool Version: Professional manual.* Lutz, FL: Psychological Assessment Resources, Inc.

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

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

Godinez, D. A., Rhee, S. H., Friedman, N., Young, S. E., Corley, R., DeFries, J. C., et al. (2005). Genetic and environmental architecture and structure of the Wisconsin Card Sorting Test. *Behavior Genetics, 35*(6), 802-803.

Goel, V. & Dolan, R J. (2004). Differential involvement in left prefrontal cortex in inductive and deductive reasoning. *Cognition, 93*, B109-B121.

Goldberg, E. & Podell, K. (2000). Adaptive decision making, ecological validity, and the frontal lobes. *Journal of Clinical and Experimental Neuropsychology, 22*(1), 56-68.

Goldberg, M. C., Mostofsky, S. H., Cutting, L. E., Mahone, E. M., Astor, B. C., Denckla, M. B., et al. (2005). Subtle executive impairment in children with autism and children with ADHD. *Journal of Autism and Developmental Disorders, 35*(3), 279-293.

Greve, K., W., Stickle, T. R., Love, J. M., Bianchini, K. J., & Stanford, M. S. (2005). Latent structure of the Wisconsin Card Sorting Test: A confirmatory factor analytic study. *Archives of Clinical Neuropsychology, 20*, 355-364.

Greve, K. W., Williams, M. C., Haas, W. G., Littell, R. R., & Reinoso, C. (1996). The role of attention in Wisconsin Card Sorting Test performance. *Archives of Clinical Neuropsychology, 11*(3), 215-222.

Guy, S. C., Isquith, P. K., & Gioia, G. A. (1996). *Behavior Rating Inventory of Executive Function—Self-Report Version: Professional manual.* Lutz, FL: Psychological Assessment Resources, Inc.

Hale, J.B., Blaine-Halpern, D., & Beakley, K. (2007, February). *Executive impairment determines ADHD response to methylphenidate treatment.* Paper presentation at the 35th Annual Meeting of the International Neuropsychological Society, Portland, OR.

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

Hale, J. B., Fiorello, C. A., & Brown, L. L. (2005). Determining medication treatment effects using teacher ratings and classroom observations of children with ADHD: Does neuropsychological impairment matter? *Educational & Child Psychology, 22*(2), 39-61.

Hale, J. B., Hoeppner, J. B., & Fiorello, C. A. (2002). Analyzing digit span components for assessment of attention processes. *Journal of Psychoeducational Assessment, 20*, 128-143.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual.* Lutz, FL: Psychological Assessment Resources, Inc.

Heinrichs, R. W. (1990). Variables associated with Wisconsin Card Sorting Test performance in neuropsychiatric patients referred for assessment. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 3*(2), 107-112.

Henry, J. D. & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. *Neuropsychology, 18*(4), 621-628.

Hirshorn, E. A. & Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: neural correlates of switching during verbal fluency. *Neuropsychologia, 44*, 2547-2557.

Holmes, A. J. & Pizzagalli, D. A. (2007). Task feedback effects on conflict monitoring and executive control: relationship to subclinical measures of depression. *Emotion, 7* (1), 68-76.

Kaplan, E. (1990). The process approach to neuropsychological assessment of psychiatric patients. *Journal of Neuropsychiatry, 2,* 72-87.

Kaplan, G. B., Sengor, N. S., Gurvit, H., Genc, I., & Guzelis, C. (2006). A composite neural network for perseveration and distractibility in the Wisconsin Card Sorting Test. *Neural Networks, 19*, 375-387.

Kimberg, D. Y. & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behavior. *Journal of Experimental Psychology, 122* (4), 411-428.

Klin, A., Volkmar, F. R., & Sparrow, S. S. (2000). *Asperger Syndrome.* New York: Guilford Press.

Konishi, S., Hayashi, T., Uchida, I., Kikyo, H., Takahashi, E., & Miyashita, Y. (2002). Hemispheric asymmetry in human lateral prefrontal cortex during cognitive set shifting. *Proceedings of the National Academy of Sciences of the United States of America,, 99* (11), 7803-7808.

Konishi, J., Kawazu, M., Uchida, I., Kikyo, H., Asakura, I., & Miyashita, Y. (1999). Contribution of working memory to transient activation in human inferior prefrontal cortex during performance of the Wisconsin card sorting test. *Cerebral Cortex, 9* (7), 745-753.

Korkman, M., Kirk, U., & Kemp, S. (1997). *NEPSY: A developmental neuropsychological assessment.* San Antonio, TX: The Psychological Corporation.

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY: A developmental neuropsychological assessment (2nd. Ed.).* San Antonio, TX: The Psychological Corporation.

Lawrence, N. S., Wooderson, S., Mataix-Cox, D., David, R., Speckens, A., & Phillips, M. L. (2006). Decision making and set shifting impairments are associated with distinct symptom dimensions in Obsessive-Compulsive Disorder. *Neuropsychology, 20*(4), 409-419.

Levine, M. D. (1999). *Developmental variations and learning disorders.* Cambridge: Educators Publishing Service, Inc.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Lichter, D. G. & Cummings, J. L. (Eds.) (2001). *Frontal-subcortical circuits in psychiatric and neurological disorders.* New York: The Guilford Press.

Lie, C., Specht, K., Marshall, J. C., & Fink, G. R. (2005). Using FMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *Journal of Neuroimage, 30,* 1038-1049.

Lopez, B. R., Lincoln, A. J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of autistic disorder. *Journal of Autism and Developmental Disorders,* 35 (4), 445-460.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology.* New York: Basic Books.

Mayberg, H. (2001). Depression and frontal-subcortical interactions. In D. G. Lichter & J. L. Cummings (Eds.), *Frontal-subcortical circuits in psychiatric and neurological disorders* (pp. 177-206). New York: Guilford Press.

McCloskey, G. (2005) *Executive functions overview.* Unpublished manuscript. Philadelphia College of Osteopathic Medicine.

McCloskey, G. (2006). *A conceptual model of executive functions.* Unpublished manuscript. Philadelphia College of Osteopathic Medicine.

McCloskey, G., & Maerlender, A. (2005). The WISC-IV Integrated. In A. Prifitera, D. H. Saklofske & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives*. Burlington, MA: Elsevier.

McCloskey, G., Perkins, L. A., & Van Divner, B. (2009). *Assessment and interventions for executive function difficulties.* New York: Routledge Press.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia output and cognition: evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition, 42*, 183-200.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (1999). The unity and diversity of executive functions and their contribution to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-100.

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., et al. (2001). Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cerebral Cortex, 11*(1), 85-92.

Nagahama, Y., Okina, T., Suzuki, N., Nabatame, H., & Matsuda, M. (2005). The cerebral correlates of different types of perseveration in the Wisconsin Card Sorting Test. *Journal of Neurology, Neurosurgery, & Psychiatry, 76*, 169-175.

Ozonoff, S. (1995). Reliability and validity of the Wisconsin card sorting test in studies of autism. *Neuropsychology,* 9 (4), 491-500.

Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R. M., Klin, A., et al. (2004). Performance on Cambridge Neuropsychological Test automated battery subtests sensitive to frontal lobe function in people with autistic disorder: Evidence from the Collaborative Programs of Excellence in Autism Network. *Journal of Autism & Developmental Disorders, 34*(2), 139-150.

Pennington, B. F. (2002). *The development of psychopathology: Nature and nurture.* New York: Guilford Press.

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

Phelps, E. A., Hyder, F., Blamire, A. M., & Shulman, R. G. (1997). FMRI of the prefrontal cortex during overt verbal fluency. *NeuroReport, 8*(2), 561-565.

Reader, M. J., Harris, E., L., Schuerholz, L. J., & Denckla, M. B. (1994). Attention deficit hyperactivity disorder and executive dysfunction. *Developmental Neuropsychology, 10*(4), 493-512.

Reynolds, C. R. & Kamphaus, R. W. (1992). *Manual for the Behavior Assessment System for Children.* Circle Pines, MN: American Guidance Service Publishing.

Reynolds, C. R. & Kamphaus, R. W. (2004). *Manual for the Behavior Assessment System for Children (2nd Ed.).* Circle Pines, MN: American Guidance Service Publishing.

Riccio, C. A., Hall, J., Morgan, A., Hynd, G. W., Gonzalez, J. J., & Marshall, R. M. (1994). Executive functioning and the Wisconsin Card Sorting Test: Relationship with behavioral ratings and cognitive ability. *Developmental Neuropsychology, 10*(3), 215-229.

Romine, C. B., Lee, D., Wolfe, M. E., Homack, S., George, C., & Riccio, C. A. (2004). Wisconsin Card Sorting Test with children: A meta-analytic study of sensitivity and specificity. *Archives of Clinical Neuropsychology, 19*, 1027-1041.

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*, 532-545.

Seidman, L. J., Faraone, S. V., Biederman, J., Weber, W., & Ouellette, C. (1997). Toward defining a neuropsychology of Attention Deficit-Hyperactivity Disorder:

Performance of children and adolescents from a large clinically referred sample. *Journal of Consulting and Clinical Psychology, 65* (1), 150-160.

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

Starkstein, S. E. & Kramer, J. (2001). The disinhibition syndrome and frontal-cortical circuits. In D. G. Lichter & J. L. Cummings (Eds.), *Frontal-subcortical circuits in psychiatric and neurological disorders* (pp. 163-176). New York: Guilford Press.

Stordal, K. I. (2006). *A study of recurrent unipolar major depression and executive functions.* Unpublished doctoral dissertation. University of Bergen, Norway.

Stratta, P., Daneluzzo, E., Prosperini, P., Bustini, M., Mattei, P., & Rossi, A. (1997). Is Wisconsin Card Sorting Test performance related to ‗working memory' capacity? *Schizophrenia Research, 27*, 11-19.

Stuss, D. T., Floden, D., Alexander, M. P., Levine, B., & Katz, D. (2001). Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia, 39*, 771-786.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., et al. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia, 38*, 388-402

Tarazi, R. A., Mahone, E. M., & Zabel, T. A. (2007). Self-care independence in children with neurological disorders: An interactional model of adaptive demands and executive dysfunction. *Rehabilitation Psychology, 52* (2), 196-205.

Torrent, C., Martinez-Aran, A., Daban, C., Sanchez-Moreno, J., Comes, M., Goikolea, J. M. et al. (2006). Cognitive impairment in Bipolar II Disorder. *The British Journal of Psychiatry, 189*(3), 254-259.

Tranel, D., Damasio, H., Denburg, N. L., & Bechara, A. (2005). Does gender play a role in functional asymmetry of ventromedial prefrontal cortex? *Brain, 128*(12), 2872-2881.

Verte, S., Geurts, H. M., Roeyers, H., Oosterlaan, J., & Sergeant, J. A. (2006). Executive functioning in children with an autism spectrum disorder: Can we differentiate within the spectrum? *Journal of Autism and Developmental Disorders, 36*(3), 351-372.

Voeller, K. K. S. (2001). Attention-Deficit/Hyperactivity Disorder as a frontal-subcortical disorder. In D. G. Lichter & J. L. Cummings (Eds.), *Frontal-subcortical circuits in psychiatric and neurological disorders* (pp. 334-371). New York: Guilford Press.

Watson, T. D. (2005). *Event-related potential correlates of cognitive operations in a modified Wisconsin Card Sorting Test.* Unpublished doctoral dissertation. Stony Brook University, New York.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition Technical Manual.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *Manual for the Wechsler Intelligence Scale for Children—Fourth Edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D., Kaplan, E., Fein, D., Morris, R., Kramer, J. H., Maerlender, A., & Delis, D. C. (2004). *The Wechsler Intelligence Scale for Children—Fourth Edition Integrated.* San Antonio, TX: The Psychological Corporation.

Weiss, L. G., Saklofske, D. H., & Prifitera, A. (2005). Interpreting the WISC-IV index scores. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives.* New York: Elsevier.

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