11-1998

Test of Variables of Attention (TOVA) Utility in Differentiating Attention Deficit/Hyperactivity Disorder Subtypes

Steven Andrew Westby

TEST OF VARIABLES OF ATTENTION (TOVA) UTILITY IN DIFFERENTIATING ATTENTION DEFICIT/HYPERACTIVITY DISORDER SUBTYPES

by

Steven Andrew Westby
Bachelor of Arts, St. John's University, 1992
Master of Arts, University of North Dakota, 1994

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota
November
1998

UMI Number: 9917893

This dissertation, submitted by Steven A. Westby, in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

(Chairperson)

This thesis meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

Dean of the Graduate School Dec 1, 1998

PERMISSION

Title	Test of Variables of Attention (TOVA) Utility in Differentiating Attention Deficit/Hyperactivity Disorder Subtypes
Department	Clinical Psychology
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it free for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised by dissertation work or, in his absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Signature [signature]

Date 11/27/1998

TABLE OF CONTENTS

LIST OF TABLES

Table Page

ABSTRACT

Research has indicated that Continuous Performance Tests (CPTs) can differentiate Attention-Deficit/Hyperactivity (ADHD) subjects from controls without any psychiatric illness. However, CPTs have neither accurately differentiated ADHD children from those with other psychiatric disorders – nor differentiated subtypes of ADHD from each other. The Test of Variables of Attention (TOVA), a new CPT, has several advantages over its predecessors which may allow the TOVA to be more effective in this differentiation process.

Data from ADHD subjects was selected from children who were administered the TOVA as part of their evaluation for ADHD at Lakeland Mental Health Center in Moorhead, MN, the Child Evaluation and Treatment Program in Grand Forks, ND, and the Behavioral Health Clinic at the St. Cloud Hospital in St. Cloud, MN. Learning Disordered subjects' data was obtained from a previous study by Clay et al. (1996). Children with no history of psychiatric illness were recruited by offering research participation credit to University of North Dakota students who agreed to have their children participate in this study.

Results were evaluated by using a group (ADHD-C, ADHD-I, Learning Disordered, and non-patient control) by TOVA quartile (1,2,3,4) mixed ANOVA on all TOVA variables (using age-corrected standard scores). In addition, I computed the Positive Predictive Power (PPP), Negative Predictive Power (NPP), and Sensitivity of the

TOVA variables in order to determine the diagnostic utilityof these measures. Finally, to test a theory that "high consistency" ADHD children might outperform controls, each group was divided into halves based upon the group's response time variability scores (by a simple median split). A group by consistency (high variability vs. low variability) ANOVA was conducted on the remaining TOVA variables (errors of omission, errors of commission, and response time). Results of the PPP/NPP analyses suggested that some TOVA variables are useful in differentiating ADHD children from non-patient controls, but not useful in differentiating ADHD from LD children. Also, TOVA data do not appear to be able to differentiate ADHD subtypes from each other. Finally, the data provided little support for the theory that a subgroup of "high consistency" ADHD children would outperform controls on other TOVA variables.

First and foremost, to my beloved wife, Meg.

Second, to the many friends and family members who have inspired and supported me through the years.

Finally, to my many professors and mentors.

My deepest gratitude.

CHAPTER I: INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD) is a psychological disorder consisting of developmentally inappropriate levels of inattention, impulsivity, and hyperactivity. In operationally defining this condition, the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (DSM-IV; APA, 1994) defines problematic *inattention* as consisting of six (or more) of the following: (1) often failing to give close attention to details or making careless mistakes in schoolwork, work or other activities; (2) often having difficulty sustaining attention in tasks or play activities; (3) often not seeming to listen when spoken to directly; (4) often not following through on instructions and failing to finish schoolwork, chores, or duties in the workplace; (5) often having difficulty organizing tasks and activities; (6) often avoiding, disliking, or being reluctant to engage in tasks requiring sustained mental effort, (7) often losing things necessary for tasks or activities; (8) often being distracted by extraneous stimuli; and (9) often being forgetful in daily activities.

The second category of ADHD symptoms presented in the DSM-IV is *hyperactivity/impulsivity*; diagnosable difficulties in this area require that six (or more) symptoms of hyperactivity and impulsivity are met. Clinically significant *hyperactivity* is defined as including the following: (1) often fidgeting with hands or feet or squirming in seat; (2) often leaving seat in classroom or other situations where remaining seated is expected; (3) often running about or climbing excessively in situations where it is inappropriate; (4) often having difficulty playing or engaging in leisure activities quietly; (5) being often "on the go" or often acting as if being "driven by a motor"; and (6) often talking excessively. Clinically significant *impulsivity* is defined by the DSM-IV as including the following: (1) often blurting out answers before questions have been completed, (2) often having difficulty awaiting their turn, and (3) often interrupting or intruding on others. In order to receive a DSM-IV diagnosis of ADHD, at least some of these symptoms must have been present before seven years of age, impairment from these symptoms must be demonstrated in at least two or more settings, and there must be clear evidence of impairment in social, academic, or occupational functioning. Although the DSM-IV dictates that there must be clear evidence of impairment in the child's functioning, precisely what constitutes "impairment" is not specified and is left up to clinical judgment.

From this grouping of symptoms, the DSM-IV divides ADHD into three discrete subtypes: a primarily inattentive subtype, a primarily hyperactive/impulsive subtype, and a combined subtype. These subtypes represent the type of ADHD symptoms that result in clinically significant impairment. Hence, the primarily inattentive subtype indicates a pattern of clinically significant difficulties with inattention (without concomitant difficulties with hyperactivity/impulsivity), whereas the combined subtype indicates clinically significant difficulties with both inattention and hyperactivity/impulsivity. This method of dividing ADHD into subtypes was first proposed in the DSM-III, but was less prominent in the DSM-III-R (APA, 1987), because of the lack of research at that time supporting the usefulness of this approach (Barkley, 1990).

The DSM-IV, however, marked a return to dividing ADHD into subtypes (i.e., by the presence or absence of hyperactivity/impulsivity behaviors) – primarily in order to make more homogenous subgroups out of a heterogeneous population. Indeed, it has been noted that ADHD children with hyperactivity and ADHD children without hyperactivity have markedly different "psychiatric symptoms, family backgrounds, developmental courses and responses to treatments" (Barkley, 1990, p. 172).

Current Conceptualizations of ADHD

Although there is widespread agreement regarding the central characteristics of ADHD, the nature of the disorder itself has remained enigmatic (Wicks-Nelson & Israel, 1997). For instance (and despite the title of the disorder), research has not consistently demonstrated a specific *attentional* deficit in children diagnosed with ADHD. For example, selective attention (the ability to attend to relevant stimuli while simultaneously ignoring irrelevant stimuli), has been theorized to be deficient in ADHD children because of their well-known behavioral tendencies to spend excessive time attending to task-irrelevant stimuli and engaging in task-irrelevant activities (Campbell & Werry, 1986). However, empirical studies using objective measures of attention have not consistently demonstrated the presence of a selective attention deficit. Some investigators have found that attention-deficit children show impaired performance when irrelevant stimuli are presented (e.g., Douglas, 1983). Other studies, however, have not demonstrated that ADHD children perform more poorly when irrelevant stimuli are presented – or have demonstrated that irrelevant stimuli impair the performance of ADHD and non-ADHD children to similar degrees (Aman & Turbott, 1986; McIntyre, Blackwell, & Denton, 1978; Radosh & Gittelman, 1981; Rosenthal & Allen, 1978, 1980). Based on these findings, several reviewers have noted that the research overall does not appear to support the absolute presence of a deficit in selective attention in ADHD children (Taylor, 1994; Whalen, 1989).

Sustained attention (the ability to pay attention to a stimulus over a specific period of time) has also been studied in ADHD children; results in this regard have not consistently demonstrated that performance by ADHD children worsens as the length of a task increases, and thus does not strongly support the existence of a deficit in sustained attention (Corkum & Seigel, 1993; van der Meere, Wekking, & Sergeant, 1991).

Because of the failure to find a specific attentional deficit, cognitively oriented researchers have been increasingly describing ADHD symptoms as resulting from deficiencies in higher-level cognitive processes (Taylor, 1994). White and Sprague (1992), for example, found that ADHD children did less planning and systematic comparison of stimuli than controls on a matching task (Matching Familiar Figures Test (MFFT); Kagan, 1964); this would appear to implicate cognitive processes involving the regulation and allocation of attention (executive functions). These findings were consistent with previous research by Chelune, Ferguson, Koon, and Dickey (1986), who compared ADHD children to controls on the Wisconsin Card Sort Test (WCST; Heating, 1981), and the Progressive Figures Test. These authors discovered that ADHD children performed more poorly than controls on these measures. In addition, Boucugnani and

Jones (1989) compared ADHD and normal controls on the WCST, Trail Making Test, and the Stroop Test – and also discovered deficits in ADHD children (relative to controls) on these measures of executive functions. Finally, support for this position was found in a study by Pennington, Groisser, and Walsh (1993). These authors compared children with reading disability (RD), children with ADHD, and co-morbid children (children diagnosed with both ADHD and RD) on two types of tasks: tasks thought to test phonological processes and tests thought to measure executive functions. Specifically, the Pennington et al. study used the Wisconsin Card Sort Test (WCST), the Tower of Hanoi task, a Continuous Performance Test, and the Matching Familiar Figures Test as measures of executive functions. In this study, RD and ADHD+RD children were shown to have deficits in phonological processing, but did not have deficits in executive functions. In contrast, ADHD children (without a co-morbid RD diagnosis) were found to have executive function deficits (as measured by these instruments). The authors argue for a "phenocopy" hypothesis – wherein ADHD+RD children display the same behavioral characteristics as ADHD children, but with a different underlying cognitive profile.

However, evidence for the executive functions deficit hypothesis has not been equivocal. Weyandt and Willis (1994), for instance, compared children with ADHD, developmental language disorder, and control children with no history of psychiatric illness on six tests of executive functions (the WCST, the Matching Familiar Figures Test, the Visual Search (Welsh, Pennington, & Groisser, 1991) test, the Verbal Fluency (Welsh et al., 1991) test, the Tower of Hanoi (Borys, Spitz, & Dorans, 1982), and the

WISC-R Mazes subtest), along with two non-executive function tasks (the Peabody Picture Vocabulary Test – Revised (PPVT-R; Dunn & Dunn, 1981) and the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983)). ADHD children were found to differ from controls on two measures of executive function (MFFT, Tower of Hanoi), but not on the Verbal Fluency, Visual Search, WCST, or mazes tasks. In addition, the ADHD group did not significantly differ on these measures from the developmental language disorder group. Thus, the authors argue that, while executive function deficits are found in ADHD subjects, these deficits do not appear to be unique to ADHD. In addition, Narhi and Ahonen (1995) compared Reading Disordered (RD), RD+ADHD, ADHD, and control children on somewhat different tests of executive functions (perseverative errors on the WCST, and the time taken to complete the Trail Making Test – Part B (TMT-B; Reitan & Wolfson, 1985)) than the Pennington et al. (1993) study. In this study, all clinical groups were found to have deficits on the measures of executive functions, and did not differ significantly from one another; thus, ADHD children were not found to be unique in this regard.

The discrepancy in research findings on the role of executive functions in ADHD has led to much speculation regarding the differences in findings. Douglas (1988), proposing that ADHD is the result of self-regulatory deficits, argued that this deficit would present itself differently depending upon reinforcement schedules and processing load. Thus, as Narhi and Ahonen (1995) suggest, it may be that the aforementioned differences reflect differences in the studies' test settings (e.g., the total number of tests used, the length of the test setting, the order in which the tests are presented) and variability in reinforcement contingencies (e.g., encouragement vs. negligence during testing). Second, the results may simply reflect the heterogeneity of ADHD children as a group. Since ADHD is typically defined by behavioral criteria alone, it is possible that "ADHD behaviors" can result from a variety of underlying factors – of which executive function deficits may be only one. If so, one would not expect to find executive function deficits consistently in ADHD children. Finally, the discrepancy may reflect the difficulty in measuring executive functions. Since "executive functions" by definition (Denckla, 1991) are controlling functions, any test which measures an "executive function" also will be influenced by deficits in more basic functions (e.g., linguistic or visual-spatial reasoning). For instance, poor performance on the Trail Making Test – Part B (TMT-B) may be due to executive function deficits, but may also reflect poorly automatized alphabets, difficulties in the visual-spatial domain, or fatigue (to name but a few). Therefore, Narhi and Ahonen's (1995) finding that all clinical groups performed worse than controls on measures of executive functions (see above) may not indicate that all clinical groups suffer from executive function deficits – but rather may suggest that different clinical groups have deficits in different areas.

In contrast to the executive functions deficit hypothesis, others have described ADHD as a deficit in motivation and behavioral regulation. Barkley (1990), for instance, has noted that in normal development behavior comes under the control of socially-relevant stimuli – such as the consequences of behavior, the requests and rules of adults, and the environmental setting. In Barkley's theory, the control of behavior by these stimuli is inadequate in ADHD children. Specifically, he argues that ADHD may stem from a diminished sensitivity to behavioral consequences, the diminished control of behavior by partial schedules of consequences, and/or poor rule-governed behavior. Barkley notes that an important reason for these deficits might be abnormally high thresholds for reinforcement. This would help to explain why ADHD children require unusually strong and salient reinforcers; it may also explain why these children fail to pay attention, comply with directions, or persist at tasks when consequences are inconsistent or weak. High thresholds for arousal could also be implicated in the child's heightened activity level and inattention. Barkley's analysis thus emphasizes biologically based deficits in the regulation of behavior by rules or consequences – rather than attention or other cognitive deficits.

Of course, both the attentional and motivational theories implicate deficits in frontal lobe functioning (Barkley, 1990). This is consistent with studies demonstrating that children with ADHD have been found to have decreased blood flow, glucose utilization, and EEG activation in the frontal lobes (Hechtman, 1991; Taylor, 1994; Zametkin & Rapoport, 1986). Neuropsychological tests (e.g., CPTs) have generally shown deficits in inhibiting motor responding in ADHD children – which also tends to suggest frontal lobe involvement (Barkley, Grodzinsky, & DuPaul, 1992).

Differences in ADHD Subtypes

Early studies, employing DSM-III terminology, compared Attention Deficit Disorder with hyperactivity (ADD/+H) to Attention Deficit Disorder without hyperactivity (ADD/-H), with mixed results. Some descriptive studies found few, if any, important differences between the two groups (Maurer & Stewart, 1980; Rubinstein &

Brown, 1984). In contrast, other studies have demonstrated that ADD/+H children displayed higher levels of aggressiveness, lower self-esteem, greater impairment on cognitive and motor tests, and were more likely to be rejected by peers than ADD/-H children (Berry, Shaywitz, & Shaywitz, 1985; King & Young, 1982). ADD/-H children, in contrast to ADD/+H children, were found to be more anxious, lethargic, sluggish, and daydreamy (Edelbrock, Costello, & Kessler, 1984; Lahey, Schaughency, Hynd, Carlson, & Nieves, 1987; Lahey, Schaughency, Strauss, & Frame, 1984).

Studies have also found mixed results in the area of academic impairment. Most studies have demonstrated no significant differences between ADD/+H and ADD/-H subjects, although both groups demonstrated significant impairment relative to controls (Barkley, DuPaul, & McMurray, 1990, 1991; Carlson, Lahey, & Neeper, 1986; Lahey, 1988). A few studies, however, have discovered a greater incidence of Learning Disorder (LD) – and thus greater academic impairment – in ADD/-H subjects (e.g., Edelbrock, Costello, & Kessler, 1984; Hynd, Lorys-Vernon, Semrud-Clikeman, Nieves, Huettner, & Lahey, 1993).

Early attempts were also made to examine possible neuropsychological differences between the two groups, and seemed to suggest that ADD/+H and ADD/-H children shared a similar profile in this regard. Carlson et al. (1986) compared ADD/+H and ADD/-H groups on the Stroop (1935) test, which is thought to measure response inhibition. No significant differences were found between groups. Using the Luria-Nebraska Neuropsychological Battery – Children's Revision (LNNB-CR), Schaughency, Lahey, Hynd, Stone, Piacentini, and Frick (1989) found no differences between ADD/+H and ADD/-H children, as well as no differences between these groups and controls.

Barkley, Grodzinsky, and DuPaul (1992) have criticized these early studies, however, on a number of grounds. First, many of these early studies relied on clinician's judgments using DSM-III criteria for placement of subjects into these various subtypes. Factor analytic studies, however, have demonstrated significant problems with using DSM-III criteria to divide subjects into subtypes. Lahey, Pelham, Schaughency, Atkins, Murphy, Hynd, Russo, Hartdagen, and Lorys-Vernon (1988), for example, conducted a factor analysis study of DSM-III ADHD criteria; they discovered that the symptoms did not cluster into the same behavioral dimensions as they are listed in the DSM-III. Specifically, DSM-III items for "impulsivity" were often strongly correlated with items for "hyperactivity" – and formed a single dimension in the factor analysis. Dividing the subjects into groups based on DSM-III criteria is likely to have provided the false impression that children deemed to have significant impulsivity were qualitatively different from those diagnosed with significant hyperactivity – whereas, in fact, the research suggests that the DSM-III criteria for impulsivity and hyperactivity essentially measure a single dimension. Thus, dividing the groups based on DSM-III criteria was likely to result in impure subgroups, making the interpretation of results difficult. A second criticism has been that several early studies used non-clinical samples of children and relied solely on teacher ratings for group placement. Since ADHD must, by definition, produce impairments in functioning across a variety of contexts, relying solely on teacher ratings (i.e., measuring impairment in only one environment) brings into question the applicability of these findings to those suffering from clinically significant conditions (i.e., those who demonstrate impairment across a variety of settings). Finally, most early studies did not employ a control group of Learning Disabled (LD) children. Twenty to fifty percent of ADD children are thought to also have co-existing LD (Barkley et al., 1990). Since academic failure is a primary characteristic of both ADHD and LD – and since ADHD and LD are significantly correlated – a "pure" LD group (i.e., a group with LD but without ADHD) is thought to be an important control group. Hence, Barkley and colleagues have argued that the degree to which the findings (or lack thereof) in these studies are due to the presence or absence of LD in the groups is uncertain, and should not be attributed to ADHD.

In addition to these general criticisms of various early studies, the weaknesses of several studies have been pointed out specifically (e.g., by Barkley et al., 1992). The Carlson et al. (1986) study – noted above – indicated no differences between ADD/+H and ADD/-H subjects on the Stroop test; this single measure, however, can hardly be thought to reflect the entirety of functions served by the frontal lobes. Schaughency et al. (1989) found no differences between ADD subtypes on the Luria-Nebraska Neuropsychological Battery – Children's Revision (LNNB-CR); however, this test was criticized as having no scales which specifically assess frontal lobe functioning – and, indeed, was designed to avoid measuring frontal lobe functions (Barkley et al., 1992, p. 174). Since current theories regarding the etiology of ADHD implicate a deficit in processes controlled by the frontal lobes, these early studies would seem not to test the appropriate areas of neuropsychological functioning.

In contrast to the above studies, several investigations have reported significant differences between ADD/+H and ADD/-H subjects. For instance, a variety of studies (e.g., Barkley et al., 1990; Lahey et al., 1988; Sergeant & Scholten, 1985) have demonstrated that the ADD subtypes may have deficits in different areas of attention. Barkley (1990), for example, demonstrated that ADD/-H children performed significantly worse than ADD/+H and control children on the Coding subtest of the WISC-R (Wechsler, 1974), while ADD/+H children did not significantly differ from controls on this measure. In contrast, ADD/+H children showed more off-task behaviors (e.g., looking away from the computer screen) than ADD/-H children during a vigilance test, and were generally described as more aggressive, impulsive, and overactive both at home and at school. In addition, research has suggested that ADD/+H children perform more poorly than ADD/-H children on the Stroop test and the Hand Movements subtest of the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983), while ADD/-H children did not differ from Learning Disabled (LD) children or controls on these measures (Barkley, 1990). Barkley has argued, based on these results, that ADD/+H and ADD/-H represent distinct disorders, with ADD/+H children primarily deficient in sustained effort during boring tasks and ADD/-H children primarily suffering from a slower perceptual-motor processing speed or impairment in focused attention. Barkley notes that this argument is consistent with the evidence that ADD/+H and ADD/-H children also show distinct patterns of familial psychiatric disturbance (Barkley et al., 1990); ADD/+H children have a greater incidence of conduct difficulties, hyperactivity, and alcohol abuse in their families, while ADD/-H children have more relatives with anxiety disorders and LDs. Again, this distinct pattern of familial psychiatric history would tend to suggest that ADD/+H and ADD/-H both represent distinct clinical disorders, rather than subtypes of a single disorder.

Barkley et al. (1992) attempted to address some of the aforementioned difficulties by comparing ADD/+H, ADD/-H, LD, and control children on a variety of neuropsychological measures thought to be sensitive to frontal lobe dysfunction – the Continuous Performance Test (Gordon, 1983), the Grooved Pegboard Test (Reitan & Wolfson, 1985), the Controlled Word Association Test (Benton & Hamsher, 1978), the Hand Movements Scale (Kaufman & Kaufman, 1983), Porteus Mazes (Porteus, 1965), Rey-Osterrieth Complex Figure (Lezak, 1983), Stroop Color-Word Test (Stroop, 1935), Trail Making Test (Reitan & Wolfson, 1985), and the Wisconsin Card Sorting Test (Heaton, 1981). Only the CPT and the Stroop were found to reliably distinguish among the groups. ADD/+H and ADD/-H subjects both made more omission errors on the CPT than the control group, and all clinical groups performed more poorly relative to controls on the Stroop Test. No significant differences were discovered between ADD/+H and ADD/-H children on these measures. Barkley et al. (1992) note that this study was limited in its relatively small sample size (approximately 12 subjects per group), and in its assumption that measures known to be sensitive to frontal lobe dysfunction in adults would also indicate such dysfunction in children. It was noted that measures sensitive to frontal lobe dysfunction in adults may not necessarily be sensitive to such dysfunction in children (Taylor, Fletcher, & Staz, 1984).

In sum, the data appear convincing that there are some significant differences between ADD/+H and ADD/-H children. These groups appear to have unique patterns of familial psychiatric disturbance. They also appear to differ in their level of aggressiveness, self-esteem, peer-rejection, and anxiety level. Studies employing neuropsychological tests have provided mixed results. Some results have suggested that these groups perform quite similarly on a variety of measures; these studies have been attacked on a variety of methodological grounds. Other studies have demonstrated that ADD/-H children have a unique difficulty with perceptual-motor speed and processing. Overall, the data appears to suggest that a unique cognitive deficit involving perceptual-motor speed and processing characterizes ADD/-H children (in contrast to ADD/+H children). Since Continuous Performance Tests (CPTs) typically measure response speed (along with impulsivity and inattention), there has been a fair amount of research interest in the use of CPTs to assess for ADHD (in general) and ADD/-H (or the current DSM-IV equivalent, Attention Deficit Hyperactivity Disorder, primarily inattentive type (ADHD-I)) in particular.

Continuous Performance Tests

In recent years, Continuous Performance Tests (CPTs) have become increasingly used as objective measures of inattention and impulsivity. A variety of CPTs have been developed, but most involve monitoring a series of stimuli for a predetermined target. Gordon (1983), for instance, developed a CPT wherein numbers are displayed on a computer screen at the rate of one per second for a period of 9 minutes; the subject is instructed to respond whenever a "9" follows a "1" (e.g., 19). Generally, CPT tests produce measures of commission (responding in the absence of the predetermined stimulus) and omission (failing to respond in the presence of the predetermined stimulus). Errors of commission have generally been thought to reflect impulsivity, whereas errors of omission have been thought to reflect inattention (Barkley, 1990). Some CPT tests also provide measures of response latency and the variability of response latency. ADHD children perform poorly relative to controls on measures of commission errors, omission errors, response latency, and the variability of response latency (Greenberg & Waldman, 1993).

Various studies have shown that ADHD subjects perform significantly worse than control subjects on CPTs do (for a review, see Barkley, 1991). However, a variety of other conditions have been demonstrated to impair CPT performance, such as children at risk for schizophrenia (Nuechterlein, 1983), learning disabled children (Dainer, Klorman, Salzman, Hess, Davidson, & Michael, 1981), and hypoxic children (O'Dougherty, Nuechterlein, & Drew, 1984). Hence, two major questions have been raised regarding CPTs: what precisely do CPTs measure, and are CPTs useful in differentially diagnosing ADHD from other clinical disorders?

Klee and Garfinkel (1983) reported significant correlations between total CPT errors and attention (measured by the Conners (1969) Teacher Rating Scale) in child psychiatric patients. Total errors on the CPT also correlated significantly with impulsivity, as measured by the Kagan (1964) Matching Figures Test. Errors of omission were found to significantly correlate with the Arithmetic subtest of the Wechsler Intelligence Scale for Children – Revised (WISC-R; Wechsler, 1974). Klee and

Garfinkle (1983) argue, based on their results, that CPT scores are most clearly related to impulsivity.

Campbell, D'Amato, Raggio, and Stevens (1991) examined the construct validity of the CPT. These authors administered the CPT, the WISC-R, the Wide Range Achievement Test – Revised (WRAT-R; Jastak & Wilson, 1984), the Bender Visual-Motor Gestalt Test (VMGT; Bender, 1938), the Conners Parent Behavior Rating Scale (Conners, 1979), and the Reading Comprehension subtest from the Peabody Individual Achievement Test – Revised (PIAT-R; Markwardt, 1989) to a group of children with learning problems. Using factor analyses, the researchers argued that CPT results were more clearly related to academic achievement than to verbal intelligence, student behavior, or perceptual-spatial organizational abilities.

Halperin, Wolf, Pasculvaca, Newcorn, Healey, O'Brien, Morganstein, and Young (1988) examined the question of what Continuous Performance Tests measure by using an A-X CPT (developed by Rosvold, Mirsky, Sarason, Bransome, Jr., & Beck, 1956). In this CPT, children are asked to press a button whenever a visually presented "X" is preceded by an "A" (e.g., AX) on a computer screen. These authors proposed that there are omission errors (not responding to an "X" when preceded by an "A") and various types of commission errors on this instrument. Specifically, they noted that a child could respond to letters other than "X" following an "A" (e.g., an "A-not-X" commission error), respond to an "X" not preceded by an "A" (e.g., an "X-only" commission error), or respond simply to an "A" (an "A-only" commission error). The subjects in this study were 85 children between first and sixth grade; these children were administered the CPT, and their teachers were asked to complete the revised Conners Teacher's Questionaire (CTQ; Goyette, Conners, & Ulrich, 1978) and a rating scale based on DSM-III criteria for ADD/+H (which measured inattention, impulsivity, and hyperactivity). Omission errors were significantly correlated with CTQ ratings of inattention, and with the DSM-III scale for inattention. "A-not-X" errors were significantly correlated with CTQ ratings of conduct problems and hyperactivity, and with DSM-III scale ratings of impulsivity and hyperactivity. "X-only" errors were found to be significantly correlated with CTQ ratings of inattention.

Lassiter, D'Amato, Raggio, Whitten, and Bardos (1994) administered a version of CPT known as the Raggio Evaluation of Attention Deficit Disorder (READD; Raggio, 1991) and a variety other measures to 104 children referred to a medical center for learning difficulties. The READD presents letter stimuli at 0.8-second intervals for a period of 8 minutes and 40 seconds. CPT scores (i.e., errors of omission and commission) on this measure were unrelated to academic functioning (as measured by WRAT-R scores), but commission errors did correlate with a teachers' reports of oppositional behavior on the ADD-H Comprehensive Teacher's Rating Scale (ACTeRS; Ullman, Sleator, & Sprague, 1986) and parents' reports of hyperactivity on the Conners Parent Behavior Rating Scale (Conners, 1979). However, CPT errors of omission were unrelated to measures of inattention on the ACTeRS. This study supports the notion that errors of commission reflect impulsivity/hyperactivity, but provides no support for the commonly held assumption that CPT errors of omission reflect inattention.

Lovejoy and Rasmussen (1990) attempted to measure the validity of vigilance tasks (including the CPT), with 100 children referred for attention and learning difficulties. These children were administered the Children's Checking Task (CCT; Margolis, 1972), a 20 minute vigilance test requiring the child to "check" discrepancies between two nearly identical series numbers in booklets, and a visual CPT developed by Lindgren and Lyons (1984). This version of the CPT lasts approximately 2.5 to 3 minutes, and the child is instructed to respond when an orange "H" on the screen precedes a blue "T". Children were also administered the Matching Familiar Figures Test (Kagan et al., 1964), wherein the children are presented with one figure and six highly similar facsimile figures and are directed to choose the variant which precisely matches the standard. Finally, children were administered the Freedom from Distractibility index subtests from the WISC-R. Parents in this study completed the Conners Parent Rating Scale and the Achenbach Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983). The children's teachers were asked to complete the Conners Teacher Rating Scale. These authors discovered significant (albeit moderate) correlations between laboratory measures of attention and impulsivity (i.e., between the CPT, CCT, MFFT, and FFD), but no correlation between vigilance task scores (i.e., scores on the CPT and CCT) and parent or teacher ratings of behavior (as measured by the Conners scales or the CBCL). The authors discovered more evidence of convergent validity for the CCT than for the CPT. The authors contend that the CCT is a longer vigilance task than their version of the CPT (lasting 20 minutes in contrast to the CPT's 2.5 to 3 minutes), and hence was more likely to uncover deficits in sustained attention.

Overall, the evidence regarding what is measured by CPTs appears mixed. Some studied have suggested that CPT variables do correlate with behavioral measures of inattention, while other studies have not supported this position. Interpretation of this set of research is complicated by various factors. First, CPTs of markedly different styles and lengths were used, making it difficult to determine to what extent the different findings resulted from the specific CPT used. Second, several different behavioral measures of inattention and hyperactivity were employed, and data regarding the correlations between these measures is not present. Finally, the CPTs employed frequently involved the presentation of numeric or alphabetic characters (e.g., "1" or "A") without employing a LD control group; since ADHD is known to be correlated with LD – and LD children are known to be slower in processing numeric/alphabetic characters – it is difficult to determine to what extent the findings may reflect the relative prevalence of LD in the groups.

The second question to be addressed is the degree to which CPTs are known to be useful in differentiating ADHD subtypes – and, more generally, ADHD from other clinical disorders. Matier-Sharma, Perachio, Newcorn, Sharma, and Halperin (1995) administered the A-X CPT with a duration of 12 minutes to ADHD subjects, non-ADHD patients, and controls. The sensitivity (i.e., the proportion of subjects with a known diagnosis who receive a positive finding on a measure) and the specificity (i.e., the proportion of subjects without a diagnosis to receive a negative finding on a measure) of CPT scores was assessed. Cut-off scores for what constituted a "positive finding" were derived by comparing subjects' scores with a normative sample; those with a score greater than 1.5 standard deviations above the mean (based on Halperin, Sharma, Greenblatt, & Schwartz, 1991) were considered to have a "positive" finding (i.e., a finding considered to be unusually discrepant from the average score in normal subjects) on the CPT. CPT errors of omission were found to have moderate sensitivity to ADHD (.70), but low specificity (.83) when comparing ADHD to controls (the specificity fell to .51 when comparing ADHD subjects to non-ADHD patients). CPT errors of commission were found to have low sensitivity (.23), but high specificity (.94) when comparing ADHD to controls (the specificity fell slightly to .88 when comparing ADHD subjects to non-ADHD patients). Classification was generally superior when comparing ADHD subjects to controls than when comparing ADHD to non-ADHD patients.

Barkley and Grodzinsky (1994) re-examined data from a previous study which compared ADD/+H, ADD/-H, non-ADD Learning Disabled (LD), and a control group on various measures (Barkley et al, 1992) to examine their Positive Predictive Power (PPP; the probability of having a condition given an abnormal score on a measure) and Negative Predictive Power (NPP; the probability of not having a condition given the absence of an abnormal score on a measure). Positive Predictive Power (PPP) and Negative Predictive Power (NPP) are constructs similar to that of sensitivity and specificity. There are important differences, however. Sensitivity and specificity are meant to address the question: given membership in a particular group, what is the probability that an individual will have a "positive" (or significantly discrepant) finding on a particular measure? In contrast, PPP and NPP examine the efficiency of a particular measure; PPP reflects the ratio of true positives to all positive findings on a measure, whereas NPP represents the ratio of true negatives to all negatives on a measure. In short, PPP and NPP address a slightly different question: given a particular finding on a test, what is the likelihood that an individual belongs in a corresponding group? As noted earlier, Gordon's (1983) CPT was used as an objective measure of attention difficulties. In this study, errors of commission had a PPP of .63, and a NPP of .82 for ADD/+H group membership. Errors of omission had a PPP of .33 and a NPP of .77 for ADD/+H group membership. When ADD/+H and ADD/-H were considered as a unitary group, commission errors improved in their PPP to 1.00, while the NPP for this measure was reduced to .59. Similarly, when errors of omission were used to classify ADD/+H and ADD/-H as a single group, PPP improved to a .92, while NPP declined to .63. Hence, abnormal scores on these measures appear to differentiate ADD subjects from control and LD subjects, but do not accurately differentiate subtypes of ADHD from each other.

In short, the literature regarding CPTs has been somewhat mixed. On the one hand, CPTs have been found to consistently distinguish ADHD from control subjects (Barkley, 1991). However, the research has been inconsistent in demonstrating the behavioral correlates (e.g., inattentiveness, oppositionality) of CPT performance, and early versions of the CPT described above have not been shown to adequately distinguish ADHD subtypes from each other. Efforts in this regard have likely been hampered by the dissimilarities (i.e., differences in stimuli presented and length) in CPTs studied. In any case, the limitations of earlier CPTs have led to the development of newer versions of this test. One such version, which appears promising, is the Test of Variables of Attention (TOVA; Greenberg & Waldman, 1993).

Test of Variables of Attention (TOVA)

Greenberg and Waldman (1993) have developed a relatively new version of a CPT. This test, called the Test of Variables of Attention (or TOVA), is distinct from previous CPTs in a number of respects. First, the TOVA is longer than previously reported CPTs, lasting approximately 23 minutes. This length may provide a better test of sustained attention (the ability to pay attention to a stimulus over a specific period of time). As noted earlier, previous research (e.g., Lovejoy & Rasmussen, 1990) has suggested that longer vigilance tasks tax individuals' sustained attention more strongly, and are therefore more likely to uncover deficits in sustained attention. Second, Greenberg and Waldman (1993) have published developmental norms for this measure, a feature largely missing in previous versions of CPTs. Specifically, 775 children aged 6-16 were divided into groups by age and presented the TOVA. This process yielded developmental norms for commission errors, omission errors, reaction time, and response variability. Third, the TOVA does not involve language or numerical processing. This feature is thought to be significant in that a number of children with Attention Deficit Disorders are known to also have learning difficulties – and language processing deficits may have impaired performance on previous versions of the CPT (Lambert & Sandoval, 1980). Instead, the TOVA involves discriminating between an upper and lower position relative to a fixed point on a computer screen; subjects are instructed to respond when the stimuli appears above the fixation point, and not to respond if the stimuli appears below the fixation point. A final distinction involves the rate of presentation for the target stimulus. The TOVA is a 23-minute test that is divided into four 5.75-minute quartiles.

The first and second quartiles of the TOVA present the target stimulus on 22.5% of the trials, a frequency similar to that of previous CPTs – and one that is thought to be a good test of inattention because the individual must be constantly alert in order to respond to the relatively infrequent target stimulus. The third and fourth quartiles, in contrast, present the target stimulus on 77.5% of the trials; this frequency of target stimulus presentation tends to produce response sets (as the subject "gets used to" pressing the button in response to multiple target stimulus presentations in a row), and is therefore thought to be a more sensitive test of the subject's ability to refrain from responding when the non-target stimulus is presented (and thus to be a better test of response inhibition and impulsivity). The TOVA provides measures of errors of commission, errors of omission, mean response time (RT), and mean RT variability (standard deviation).

No published studies have yet used the TOVA in differentially diagnosing subtypes of ADHD – perhaps because the TOVA is a relatively new measure, and there are few studies overall which have evaluated the differences between these subtypes. Matier-Sharma et al. (1995) report that preliminary studies comparing ADHD subjects to controls on the TOVA have yielded a sensitivity index of .68 and a specificity index of .85 in detecting ADHD when compared to non-patient controls. The TOVA's ability to differentiate ADHD from non-ADHD patients was not tested. Clay, Petros, Searcy, and Westby (1996) administered a number of psychological measures (including the TOVA) to groups of ADHD-only, LD-only, ADHD with LD, and control children. The group sizes were somewhat small (ranging from 15 in the ADHD-only group to 23 in the

ADHD+LD group), and PPP and NPP were not conducted as part of this study. However, TOVA data indicated a marginal main effect of group (i.e., the presence or absence of ADHD) on errors of commission. Errors of omission revealed a significant main effect of group (for both ADHD and LD); ADHD subjects performed significantly worse on this measure than non-ADHD children, and LD subjects performed significantly worse than non-LD children. Thus, both LD and ADHD appear to significantly affect errors of omission on the TOVA. Interestingly, a significant main effect of quartile was observed on the errors of omission variable – as was a significant ADHD by quartile interaction. Subsequent analysis indicated a sharply increasing discrepancy between ADHD and non-ADHD subjects in the third and fourth quartiles. The response time measure revealed a significant main effect of group (with LD children performing significantly worse than non-LD children) and a significant interaction of LD by quartile. Subsequent analyses suggested significant differences between LD and non-LD subjects at each quartile, but the discrepancy decreased across quartiles. Thus, while LD children were slower overall than non-LD children, the discrepancy of reaction time scores between RD and controls decreased over the length of the test – with the RD subjects showing faster response times as the test progressed. The response time variability measure (which literally measures the statistical variability of response times over the length of the test, or the standard deviation) yielded significant main effects of both ADHD and LD. In sum, the study suggests that the response time, errors of commission, errors of omission, and response time variability measures on the TOVA hold some promise in detecting the presence of ADHD and potentially differentiating ADHD children from those diagnosed with other conditions. In addition, this study appears to corroborate the usefulness of the length of the TOVA – since important information appears to be evident from the pattern of scores across quartiles.

The TOVA would appear to hold promise for assisting in the differential diagnosis of subtypes of ADHD, despite the failure of previous versions of the CPT to do so. First of all, the length of the vigilance task has been noted to be a potentially significant factor in the discrimination ability of CPTs, perhaps because the children's sustained attention is taxed (Lovejoy & Rasmussen, 1993). This notion was supported by the significant interactions of errors of omission by quartile (with the discrepancy between ADHD and non-ADHD subjects greatly increasing in the third and fourth quartiles) in the study by Clay et al. (1996). Second, as noted earlier, TOVA results may be less affected by LD than previous versions of the CPT because little language processing is necessary (Greenberg & Waldman, 1993). This notion is given partial support by the lack of LD effects on errors of commission in the Clay et al. (1996) study. Given the concerns noted by Barkley (1990) and others regarding the possible confounding of LD with ADHD in this area of research, the TOVA would appear to have a distinct advantage over previous versions of the CPT.

The presence of a response time variability measure on the TOVA provides an opportunity for an interesting analysis. In studying the cognitive effects of aging, Ferraro and Moody (1996) compared young adults and elderly adults on a measure of simple reaction time (SRT) and a measure of choice reaction time (CRT). These measures produced both mean reaction times and standard deviations of reaction times. Previous work in this area (e.g., Cerella, 1994; Fisk & Fisher, 1994) had suggested that aging produces a general slowing in cognitive functions. By dividing the elderly into two groups using a median split based on their consistency of responding (i.e., high standard deviations vs. low standard deviations on response time performance), however, Ferraro and Moody discovered that the high consistency elderly adult group (i.e., with low standard deviations) actually outperformed the young adult group. That is, the high consistency elderly were faster and more accurate than the young adult group. In relation to the TOVA, it is certainly possible that ADHD children might similarly differ in regards to their consistency of responding. If so, a "high consistency" ADHD group might reflect a subset of relatively high functioning ADHD children which could outperform controls on some other TOVA measures.

Clinical Implications

ADHD is among the more prevalent childhood disorders, frequently estimated to affect between three and five percent of the school-aged population (APA, 1994). Yet many symptoms of ADHD overlap with symptoms from other psychiatric conditions (e.g., agitated depression), thus making the differential diagnosis of this disorder difficult (Barkley, 1990). The differentiation of ADHD from other psychiatric conditions has important implications for psychological and psychiatric treatment strategy. Some evidence (reviewed earlier) also suggests that ADHD "subtypes" may reflect distinct disorders, with (for example) ADHD-I children displaying higher levels of anxiety, and ADHD-C children demonstrating more aggressiveness (Barkley et al., 1992). Thus, determining a child's ADHD subtype could also have important treatment implications.

The TOVA appears promising in its ability to differentiate ADHD from non-patient controls, and differentiating ADHD subtypes from each other. If so, the test would have a high level of diagnostic utility.

Hypotheses and Predictions

This study will investigate the ability of the TOVA to differentiate Attention Deficit Disorder, primarily inattentive type (ADHD-I), Attention Deficit Disorder, combined type (ADHD-C), learning disordered controls (LD) and non-patient control children by their performance on this measure. It is predicted that ADHD-I and ADHD-C children will perform significantly worse than controls on all TOVA measurement variables. In addition, ADHD-C children are expected to display significantly more errors of commission than ADHD-I subjects on the TOVA, while ADHD-I children will display significantly more errors of omission, slower response times, and more RT variability than ADHD-C subjects. Finally, it is predicted that "high consistency" ADHD children (i.e., children with low variability scores) will perform better than "low consistency" ADHD children and controls on the errors of omission, errors of commission, and response time measures.

CHAPTER II: METHOD

Subjects

Participating in the study were 86 children, who comprised the four groups in this study: 31 children diagnosed with ADHD-C; 15 children diagnosed with ADHD-I; 20 children diagnosed with RD; and 20 non-patient control children. All groups had children ranging in age from 7 to 12 (please see Table 2 for a breakdown of subjects by age within the groups). Please refer to Table 1 for additional demographic data (i.e., gender composition, average age, and number of subjects) regarding each group.

TOVA data was obtained from three sources. Archival TOVA data from ADHD subjects were obtained from children who had already been given the TOVA as part of their evaluation for ADHD at Lakeland Mental Health Center in Moorhead, MN, the Child Evaluation and Treatment Program in Grand Forks, ND, and the Behavioral Health Clinic at the St. Cloud Hospital in St. Cloud, MN. Subject names were taken from TOVA data files at each agency. The file of each child from this search was examined by the author to determine the appropriateness of the child for this study. Specifically, a child's data was only used in this study if the child carried a diagnosis of ADHD (combined or primarily inattentive subtype) from a licensed practitioner, and the child's diagnostic evaluation provided documentation of sufficient DSM-IV criteria to support a diagnosis of ADHD (as defined by the DSM-IV, at least six symptoms of inattention and/or hyperactivity/impulsivity must be reported for a diagnosis of ADHD to be given).

LD control subjects' TOVA data was taken from the Clay et al. (1996) study. Finally, children who have no history of psychiatric illness were be recruited from the University of North Dakota. Specifically, University of North Dakota students who were enrolled in undergraduate Psychology classes were offered research participation credit for agreeing to have their child participate in this study. In addition, the children were paid $5 for their time. Participation in this study will be limited to those children between the ages of seven and twelve years at the time of their evaluation. Approval from the Institutional Review Board (IRB) of the University of North Dakota, the Child Evaluation and Treatment Program, and the St. Cloud Hospital for the use of data in their possession was secured prior to any use of the data or TOVA administrations.

Materials

The TOVA is a 23 minute fixed-interval visual CPT which was administered on an IBM-compatible computer. Subjects are informed to respond by pressing a button (with a finger or thumb of their dominant hand) when a stimulus appears above a fixation point on a computer screen, but not to respond when the stimulus appears below the fixation point. The stimulus itself is a small, black square located either above or below the mid-line of a larger, orange square in the approximate center of the computer screen. The stimulus is randomly presented for 100 milliseconds every two seconds. Two of the quartiles (the first and second) present the correct stimulus on 22.5% of the trials. This rate of presentation is similar to previous CPTs, and is thought to be effective in detecting difficulties with inattention (Greenberg & Waldman, 1993). The third and fourth quartiles present the correct stimuli on 77.5% of the trials. This frequency was thought to be a better test of response inhibition and impulse control by being more likely to induce a response set (Greenberg & Waldman, 1993).

<u>Procedure</u>

Subjects' TOVA data will be assigned into their groups based on their DSM-IV diagnosis (or lack thereof). That is, TOVA data from children who have been diagnosed with ADHD (combined type) will be placed into the ADHD-C group, whereas the TOVA data from children given the DSM-IV diagnosis of ADHD (primarily inattentive type) will be placed in the ADHD-I group. TOVA data from children with diagnosed LD (from the Clay et al., 1996 study) will be placed into the LD group. Children were included in the Clay et al. (1996) study LD group only if they had been diagnosed with a learning disability by the local school system, they obtained a standard score less that 90 on the Word Attack subtest from the Woodcock-Johnson Psychoeducational Battery - revised (Woodcock & Johnson, 1990), and they achieved a WISC-III Performance IQ of higher than 80.

Any diagnosis of ADHD will have been arrived at by psychologists or psychiatrists at LMHC, CETP, and St. Cloud Hospital, using standard assessment procedures. The assessment procedures used by each child's psychologist or psychiatrist in arriving at the diagnoses are likely to have varied considerably across subjects. Each subject would have been given a formal diagnostic interview – and must have been administered the TOVA in order to be included in this study. The use of additional assessment measures, however, was not tracked as part of this study. Nevertheless, it is likely that some of the psychologists and/or psychiatrists employed additional methods to diagnose the children. These methods included the behavioral observation of the child, interviews with the child's teachers, additional psychological testing (e.g., WISC-III, etc.), and/or behavioral checklists completed by the parents, teachers, and other caregivers. Since this portion of the study (i.e., the placement of subjects into the ADHD subtype groups) is based upon archival data, these diagnoses were arrived at prior to the onset of this study.

As noted earlier, the non-patient (and non-archival) control group consists of children who were recruited to participate in this study. Children who agreed (and had their parents' consent) to participate in this study were asked to complete the TOVA either after school or on a weekend, and to bring a parent or guardian along. The parent was asked to complete a brief screening measure, wherein basic demographic data (e.g., date of birth, gender) was obtained. This screening measure also asked the parent about any history of psychiatric illness (including learning difficulties) in the child – and had them place a check mark next to any of the DSM-IV symptoms of ADHD (noted above) they have had observed in their child. Should the parents' report have suggested the possible presence of ADHD or another clinical disorder, the child was not be assessed using the TOVA – and a recommendation would have been issued that the parent consider having the child evaluated by a local mental health professional. There were no instances in which this was necessary during this study. If the parent's report did not suggest the presence of ADHD or another psychiatric illness, then the child was administered the TOVA. Following the TOVA administration, the child was reimbursed $5, and any questions were addressed (note: the children were reimbursed $5 even if they failed to meet the criteria for this study, and are therefore were not administered the TOVA).

As noted above, subjects in the ADHD-C, ADHD-I, and non-patient control groups were evaluated (either by chart-review or checklist) to determine the number of DSM-IV criteria for ADHD that were being observed by the child's parents. Because the LD subjects came from another study, it was not possible to conduct a similar procedure with this group. However, the LD subjects in the Clay et al. (1996) study were screened for the presence of ADHD. This was accomplished by only allowing children diagnosed with LD through the system into this group (i.e., children were not included into that study's LD group if there was an indication of ADHD in their evaluation). In any case, demographic data and the results of the present study's screening process are summarized in Table 1:

Table 1

Demographic Data and Mean Number of DSM-IV Symptoms of ADHD Reported by Chart-Review or Behavior Checklist

Group	N	Males	Females	Avg. Age	Mean # of Symptoms
ADHD-C	31	21	10	9.0	13.4
ADHD-I	15	12	3	9.6	8.1
LD	20	13	7	10.3	Not Assessed
Controls	20	15	5	9.6	0.4
Total	**86**	**61**	**25**	**9.63**	**7.30**

The number of subjects at each age (by group) is represented in Table 2:

Table 2

Age Breakdown of Subjects (as a Function of Group)

Age	ADHD-C	ADHD-I	LD	Controls	All Groups
7 years	7	4	1	2	14
8 years	5	1	1	3	10
9 years	7	1	2	6	16
10 years	7	3	6	2	18
11 years	2	3	7	4	16
12 years	3	3	3	3	12

CHAPTER III: RESULTS

Demographic Variables

A one-way ANOVA was conducted across groups to determine if the groups significantly differed by age. Results indicated no significant difference on age, $F(3,82)=2.56, p.>.05$. Similarly, a chi-square analysis was conducted on gender across groups to determine if the gender composition of the groups was significantly different. Results failed to confirm this possibility, $X^2(3)=1.29, p.>.05$.

TOVA Variables

A group (ADHD-C, ADHD-I, LD, non-patient control) by quartile (1-4) mixed ANOVA was conducted on all TOVA variables (i.e., errors of omission, errors of commission, response time, and response time variability).

An analysis of the errors of omission data suggested a main effect of quartile, $F(3,246)=5.05$, $p.<.01$, but not of group, $F(3,82)=1.84, p>.05$. No interaction was noted between group and quartile, $F(9,246)=0.93, p.>.05$. Since this analysis involved a repeated measure, Mauchly's test for the violation of the sphericity assumption was performed. Kirk (1982) explains the concept of sphericity as follows: "…a matrix whose diagonal elements are equal, and whose non-diagonal elements are zero, are said to be spherical." Violations of this assumption are known to result in increases in Type I error rates. In any case, Mauchly's test of sphericity for the quartile effect suggests that this assumption was violated, $X^2(5)=27.2$, $p<.01$; however, this effect was robust under Huynh-Feldt's adjusted ANOVA, $F(2.6, 213.9)=5.05$, $p<.01$. Effect size tests indicated a small effect of quartile, with an estimated omega squared of .05. Estimates of the (non-significant) group and group-by-quartile interaction effects also indicate small effects, with omega squares of .03 and .001, respectively. Tukey analyses indicated that, across groups, children committed fewer omission errors in the first quartile (mean score: 91.03) than in the second quartile (mean score: 83.70, $p<.05$), third quartile (mean score: 83.00, $p<.05$), or fourth quartile (mean score: 81.04, $p<.01$). Mean omission error values and standard deviations (by group and quartile) are presented in Table 3:

Table 3

Mean Omission Error Standard Scores (with Standard Deviations) as a Function of Group and Quartile

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
ADHD-C	86.35	76.41	78.54	71.32
	(23.57)	(26.44)	(28.04)	(30.05)
ADHD-I	88.53	83.87	73.20	79.80
	(24.31)	(30.27)	(30.24)	(26.58)
LD	92.60	86.80	90.35	90.50
	(24.58)	(27.33)	(21.59)	(24.79)
Controls	96.65	87.70	90.15	82.55
	(15.93)	(25.55)	(24.09)	(26.60)

Regarding errors of commission, no effect of group, $F(3,82)=1.35, p>.05$, was noted. Effect-size estimates on this variable suggest a small effect, with an estimated omega squared of .02. A main effect of quartile, $F(3,246)=3.58, p<.05$, was observed. Mauchly's test of sphericity suggests that this assumption was violated, $X^2(5)=81.25$, $p<.01$; however, this effect remained robust under Huynh-Feldt's adjusted ANOVA, $F(2.02,166.12)=3.58$, $p<.05$. Effect size estimations indicated a relatively small effect, with an estimated omega squared of .03. Tukey analyses revealed that, across groups, subjects committed fewer commission errors in the fourth quartile (mean score: 99.70) than in the second quartile (mean score: 92.47, $p<.01$). Further, a significant interaction between group and quartile was noted, $F(9,246)=2.78, p<.05$. Unsurprisingly, Mauchly's test of sphericity again suggests that this assumption was violated, $X^2(5)=27.2, p<.01$; however, this effect remained robust under Huynh-Feldt's adjusted ANOVA, $F(6.07,166.12)=2.78, p<.05$. Effect size estimations indicated a moderate effect, with an estimated omega squared of .06. Subsequent Tukey analyses suggested that ADHD-C children in the second quartile (mean score: 81.03) performed significantly worse than LD children in the second quartile (mean score: 103.30, $p<.01$). In addition, ACHD-C performed worse in the second quartile than in the third (mean score: 96.16, $p<.01$) or fourth quartiles (mean score: 101.71, $p<.01$). Mean commission error values and standard deviations (by group and quartile) are summarized in Table 4.

For the response time variable, no effect was noted for group, $F(3,82)=2.22$, $p>.05$, or quartile, $F(3,246)=0.70, p>.05$. Further, no interaction between group and quartile was noted, $F(9,246)=1.60, p>.05$. Effect-size estimates suggest a small to

Table 4

Mean Commission Error Standard Scores (with Standard Deviations) as a Function of Group and Quartile

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
ADHD-C	90.61	81.03	96.16	101.71
	(21.78)	(32.37)	(15.04)	(13.79)
ADHD-I	102.07	92.80	96.27	99.80
	(22.06)	(22.55)	(17.46)	(16.78)
LD	102.60	103.30	97.35	98.60
	(12.64)	(13.37	(12.52)	(14.28)
Controls	96.00	92.75	95.95	98.70
	(22.20)	(21.85)	(16.33)	(14.17)

moderate effect for group (estimated omega squared: .05), an extremely small effect of quartile (estimated omega squared: .0003), and a small effect of the group by quartile interaction (estimated omega squared: .02). Mean response time standard score values and standard deviations (by group and quartile) are summarized in Table 5.

On the response time variability variable, a main effect was noted for group, $F(3,82)=5.17, p.<.01$, and quartile, $F(3,246)=7.41, p.<.01$. No interaction between group and quartile was indicated, $F(9,246)=1.29, p.>.05$. Subsequent Tukey analyses indicated that, across quartiles, ADHD-C subjects (mean score: 69.33) had more variable response times than LD children (mean score: 82.56, $p.<.05$) or control children (mean score:

Table 5

Mean Response Time Standard Scores (with Standard Deviations) as a Function of Group and Quartile

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
ADHD-C	79.29	72.38	80.80	73.70
	(27.13)	(19.72)	(24.23)	(17.47)
ADHD-I	70.73	74.27	71.73	73.47
	(20.29)	(18.31)	(21.18)	(18.52)
LD	68.90	75.25	76.75	75.75
	(24.56)	(18.99)	(16.68)	(18.42)
Controls	86.80	82.75	85.30	87.45
	(16.59)	(15.75)	(17.65)	(14.41)

84.60, $p < .05$). Estimates of effect size suggest a moderately large effect, with an estimated omega squared of .14. Regarding the significant effect of quartile, Mauchly's test of sphericity suggests that this assumption was violated, $X^2(5)=17.96, p < .01$; however, this effect remained robust under Huynh-Feldt's adjusted ANOVA, $F(2.84,233.23)=7.41, p < .01$. Effect size estimations indicated a moderate effect, with an estimated omega squared of .08. In any case, Tukey analyses revealed that children across groups had less variable response times in the first quartile (average standard score: 83.92) than in the third (73.21, $p < .01$) or fourth (74.33, $p < .01$) quartiles. Effect size estimates on the (non-significant) effect of group by quartile indicate a small effect, with an estimated omega squared of .01. Mean response time variability values and standard deviations (by group and quartile) are summarized in Table 6:

Table 6

Mean Response Time Variability Standard Scores (with Standard Deviations) as a Function of Group and Quartile

	Quartile 1	Quartile 2	Quartile 3	Quartile 4
ADHD-C	78.45	66.83	68.83	63.19
	(21.09)	(23.86)	(15.28)	(21.61)
ADHD-I	79.53	77.93	68.87	71.13
	(24.36)	(23.27)	(20.37)	(21.32)
LD	85.95	89.85	73.60	80.85
	(26.14)	(20.86)	(22.03)	(22.12)
Controls	91.75	82.95	81.55	82.15
	(14.56)	(18.45)	(17.13)	(20.43)

Positive Predictive Power, Negative Predictive Power, and Sensitivity

Positive Predictive Power (PPP) and Negative Predictive Power (NPP) were computed for the four primary TOVA variables. As discussed earlier, PPP values reflect the ratio of "true positives" to all positives on a variable (and thus reflect the ability to the test to correctly "rule in" those diagnosed with ADHD); in contrast, NPP values reflect the ratio of "true negatives" to all test negatives (and thus reflect the ability of the test to correctly "rule-out" those not diagnosed with ADHD). A finding on a variable was considered to be "positive" if it fell 1.5 standard deviations (or more) from the mean (based upon norms developed by Greenberg & Waldman, 1993), and "negative" if it was not 1.5 standard deviations (or more) from the mean. A "true positive," then, would indicate a child who was diagnosed with ADHD – and received a "positive" finding on a measure. A "true negative" would indicate a child who was *not* diagnosed with ADHD that received a "negative" finding on a measure. Thus, for example, the omission errors variable had 27 "positive" findings (individuals whose standard scores were 77 or less), 18 of whom were "true positives" (i.e., had been diagnosed with ADHD). Therefore, the Positive Predictive Power (PPP) of this variable was 18/27 – or .67. In contrast, 59 "negative" findings occurred (i.e., 59 individuals had standard score greater than 77), of whom 31 were "true negatives" (i.e., they did not carry diagnoses of ADHD). Thus, the Negative Predictive Power (NPP) of this variable was 31/59 – or .53. In theory, PPP and NPP values can range from 0 (indicating no classification ability) to 1.0 (indicating perfect classification).

Errors of omission were found to have a PPP of .44 and a NPP of .68 for ADHD-C group membership (i.e., when ADHD-I subjects were not considered "true positives"). When ADHD-C and ADHD-I were considered a single group (i.e., both ADHD-C and ADHD-I subjects were considered "true positives"), the PPP for this variable improved to .67, while the NPP declined to .53. The data contributing to this analysis are represented in Table 7.

Table 7

Classification Utility Data: Omission Errors (ADHD-C & ADHD-I combined)

		TOVA Prediction		
		ADHD	Not ADHD	TOTALS
Diagnosis	ADHD:	18	28	**46**
	Not ADHD:	9	31	**40**
	TOTALS:	**27**	**59**	

Errors of commission were discovered to have a PPP of .54 and a NPP of .67 for ADHD-C group membership; the PPP improved to .77 and NPP dropped to .51 when ADHD-C and ADHD-I were combined. Data contributing to this analysis are represented in Table 8.

Table 8

Classification Utility Data: Commission Errors (ADHD-C & ADHD-I combined)

		TOVA Prediction		
		ADHD	Not ADHD	TOTALS
Diagnosis	ADHD:	10	36	**46**
	Not ADHD:	3	37	**40**
	TOTALS:	**13**	**73**	

Response times were found to have a PPP of .45 and a NPP of .73 for ADHD-C group membership. The PPP changed to .67 and the NPP to .59 when ADHD-C and

ADHD-I were combined into a single group. Data involved in this analysis are represented in Table 9.

Table 9

Classification Utility Data: Response Time (ADHD-C & ADHD-I combined)

		TOVA Prediction		
		ADHD	Not ADHD	**TOTALS**
Diagnosis	ADHD:	28	18	**46**
	Not ADHD:	14	26	**40**
	TOTALS	**42**	**44**	

Finally, response time variability was found to have a PPP of .53 and a NPP of .89 for ADHD-C group membership; the figures changed to a .71 (PPP) and .71 (NPP) when ADHD-C and ADHD-I children were combined into a single group. The data contributing to these analyses are represented in Table 10.

Table 10

Classification Utility Data: Response Time Variability (ADHD-C & ADHD-I combined)

		TOVA Prediction		
		ADHD	Not ADHD	**TOTALS**
Diagnosis	ADHD:	36	10	**46**
	Not ADHD:	15	25	**40**
	TOTALS:	**51**	**35**	

Overall, PPP and NPP appeared to improve somewhat when the LD group was not included in the classifications. For omission errors, PPP improved to .55 and NPP declined slightly to .57 (when classifying solely for ADHD-C group membership); PPP improved to .82 and NPP fell to .36 when considering ADHD-C and ADHD-I children as a single group. Based on commission errors, PPP improved to .58 and NPP fell to .56 when classifying for ADHD-C group membership; PPP improved to .83 and NPP fell to .33 when combining the ADHD groups. For response times, PPP improved to .58 while NPP fell to .56 when classifying ADHD-C alone; PPP improved to .83 and NPP declined somewhat to .33 when considering ADHD-C and ADHD-I jointly. Finally, response time variability improved its PPP to .66 while NPP fell to .84 when classifying solely the ADHD-C group; PPP was .88 and NPP .60 when considering ADHD-C and ADHD-I children together. These results are summarized in Table 11.

The sensitivity of the major TOVA variables was also computed. In contrast to PPP/NPP (which measures the percentage of those with a finding who are in a group), sensitivity measures the percentage of those with a known diagnosis (in this case, ADHD-C and/or ADHD-I) that receive a "positive" finding on a measure. Omission errors had a sensitivity of .39 for ADHD-C and of .40 for ADHD-I (with an overall sensitivity of .39 when the two subtypes are grouped together). Commission errors revealed a sensitivity of .22 for ADHD-C and of .20 for ADHD-I, with an overall sensitivity of .22 for ADHD. Mean response times had a sensitivity of .61 for ADHD-C, .60 for ADHD-I, and an overall sensitivity of .61 for ADHD. Mean response time variability (i.e., response time

Table 11

Positive Predictive Power and Negative Predictive Power for Major TOVA Variables when Classifying for ADHD-C (only) and for Combined ADHD Subtypes

	ADHD-C Classification		Combined ADHD Classification	
	PPP	NPP	PPP	NPP
LD Group Included				
Omission Errors	.44	.68	.67	.53
Commission Errors	.54	.67	.77	.51
Response Time	.45	.73	.67	.59
RT Variability	.53	.89	.71	.71
LD Group Excluded				
Omission Errors	.55	.57	.82	.36
Commission Errors	.58	.56	.83	.33
Response Time	.58	.56	.83	.33
RT Variability	.66	.84	.88	.60

standard deviations) had a sensitivity of .87 for ADHD-C, .60 for ADHD-I, and an overall sensitivity of .78 for ADHD. These results are summarized in Table 12.

Table 12

Sensitivity of the Major TOVA Variables for ADHD-C, ADHD-I, and Combined ADHD Groups

	ADHD-C	ADHD-I	Combined Groups
Omission Errors	.39	.40	.39
Commission Errors	.22	.20	.22
Response Time	.61	.60	.61
RT Variability	.87	.60	.78

Consistency

Finally, subjects' performance was divided in half by median split based upon the members' response time variability standard score (averaged across quartiles). A group (ADHD-C, ADHD-I, LD, non-patient controls) by consistency (high response time variability, low response time variability) by quartile (1-4) mixed ANOVA was conducted on the remaining TOVA variables (omission errors, commission errors, and response time). This analysis was done to test the hypothesis that "high consistency" ADHD children (i.e., children with low variability scores) would perform better than "low consistency" ADHD children and controls on the remaining major TOVA variables. On the omission errors variable, no main effect of group, $F(3,74)=0.51, p>.05$, was indicated. Main effects of consistency, $F(1,74)=8.10, p<.01$, and quartile, $F(3,222)=4.52, p<.01$, were noted however. Subsequent Tukey analyses indicated that "high consistency" children had significantly fewer omission errors (average standard score: 91.31) than "low consistency" children (average standard score: 76.79, p.<.05). Further, children across groups had fewer omission errors in the first quartile (average standard score: 90.67) than in the third (82.57, p.<.05), or fourth quartiles (79.35, p.<.05). No interactions were noted between group and consistency, $F(3,74)=1.18, p>.05$, group and quartile, $F(9,222)=0.80, p>.05$, consistency and quartile, $F(3,222)=1.12, p>.05$, or group, consistency, and quartile, $F(9,222)=0.29$, $p>.05$.

Errors of commission revealed no main effects of group, $F(3,74)=0.57, p>.05$, or quartile, $F(3,222)=2.44, p>.05$. A main effect of consistency was observed, $F(1,74)=4.22, p<.05$; subsequent Tukey analyses suggested that "high consistency" children (average standard score: 100.55) had significantly fewer commission errors than "low consistency" children (average standard score: 93.02, p.<.05). An interaction was noted between group and quartile, $F(9,222)=2.00, p<.05$. Tukey results suggested that ADHD-C children in the second quartile (average standard score: 84.51) performed significantly worse than the following: (1) LD children in the second quartile (average standard score: 103.72, p.<.01); (3) ADHD-C children in the third quartile (average standard score: 97.98, p.<.01); and (4) ADHD-C children in the fourth quartile (average standard score: 102.87, p.<.01). No interactions were noted between group and consistency, $F(3,74)=0.24, p>.05$, or consistency and quartile, $F(3,222)=0.82, p>.05$. Further, no interactions were found between group, consistency, and quartile, $F(9,222)=0.55, p>.05$.

Finally, an analysis of the response time data revealed no main effects of group, $F(3,74)=0.45, p>.05$, or quartile, $F(3,222)=0.47, p>.05$. There was, however, a main effect of consistency, $F(1,74)=10.00, p<.01$. "Low consistency" children had significantly slower response times (average standard score: 68.11) than "high consistency" children (average standard score: 80.92, $p<.01$). No interactions were noted between group and consistency, $F(3,74)=2.00$, group and quartile, $F(9,222)=1.52, p>.05$, consistency and quartile, $F(3,222)=1.60, p>.05$, or group, consistency, and quartile, $F(9,222)=1.55, p>.05$.

CHAPTER IV: DISCUSSION

This study attempted to determine the utility of the TOVA in differentially diagnosing ADHD subtypes from each other – as well as differentiating ADHD from non-ADHD children (LD children and non-patient controls). This was done by collecting archival TOVA data on ADHD subjects from three sources (see Methods), using LD data from a previous study (Clay et al., 1996), and collecting non-archival data from non-patient control children. These data were then analyzed using ANOVA, and the PPP/NPP of each variable was tabulated.

The results of this study appear to only partially support the hypothesis that ADHD-I and ADHD-C children would perform worse than controls on all major TOVA variables. In support of this hypothesis are the following findings: (1) ADHD-C children performed significantly worse than LD and non-patient control children on the response time variability measure; and (2) ADHD-C children committed significantly more commission errors than LD children *in the second quartile*. However, the results of this study failed to suggest that ADHD children (either Inattentive or Combined type) performed worse than LD or non-patient control children on the omission errors or response time variables. In addition, the ADHD-I group did not significantly differ from the LD or non-patient control groups on any of the TOVA variables. Thus, the results of this study (which may have significant limitations, discussed below) tend to suggest that the TOVA is not useful in statistically differentiating these groups.

These results are somewhat surprising, given previous research (e.g., Greenberg & Waldman, 1993) which has suggested that the TOVA variables are generally effective in differentiating ADHD subjects from non-patient control subjects. The results of the present study appear to reflect, at least in part, *unusually poor performance on many of the TOVA variables by the non-patient control subjects*. For example, the non-patient control subjects mean performance on the errors of omission variable was a standard score of 89 – almost a full standard deviation below the mean for non-patient controls developed by Greenberg and Waldman (1993). In addition, the mean response time standard score for the non-patient control group was an 86, once again almost a full standard deviation below the mean. This poor performance is somewhat difficult to explain. However, three possible explanations present themselves. First, it is possible that this study's relatively low sample-size produced these findings. To test this possibility, a power analysis was conducted using Kraemer and Thiemann's (1987) approach. Using this method, a "critical effect size" for a test (defined as "the minimum effect considered important to detect", Kraemer & Thiemann, 1987) was computed, following which the power of a study can be estimated by comparing the number of subjects per cell (needed for various levels of power) to the actual number of subjects per cell in a study. The result of this analysis suggested that this study had only a 60% chance of detecting an effect of one standard deviation or higher. Thus, there remains a 40% chance that true effects were not uncovered by this study. The power of this study appears to have been limited by two factors: a relatively small sample size (particularly in the ADHD-I group), and the discrepancy in the number of subjects per group (e.g., 31 in the ADHD-C group versus 15 in the ADHD-I group). The number of subjects used in this study is similar to that of previous research in this area (e.g., Barkley & Grodzinsky, 1994). Nevertheless, additional research in this area may wish to address these shortcomings by including more subjects – as well as attempting to make the number of subjects per group approximately equal. In favor of this argument are the results of the effect-size estimations on the non-significant effects, which suggested that some small-to-moderate effects (e.g., the estimated omega-squared of .05 for group main effects on the response time variable) may exist which were uncovered by the present study's statistical analyses.

Second, it is possible that this study had a selection bias – that is, by advertising itself as a study of attention, it is possible that parents with children who were concerned about attentional functioning may have presented them for this study. This possibility seems somewhat less likely, however, given that these parents endorsed no symptoms of ADHD. Nevertheless, it is certainly true that selection procedures for control subjects differed in this study (which relied on parents to volunteer their children) from the Waldman and Greenberg (1993) normative study, which randomly selected children before requesting that parents allow them to participate. This study's recruitment method may, therefore, have been more susceptible to a selection bias.

Third, it is possible that the children who were administered the TOVA as part of this study were given the test in a slightly different manner. Since this author was involved in both the present study and the Clay et al. (1996) study, it is reasonable to assume that the TOVA was administered similarly for these groups. However, the

ADHD-C and ADHD-I group children may have been administered the TOVA in differing ways. For example, it is possible that the subjects run as part of this study (or in the Clay et al. study) were scrutinized more (or less) closely than the ADHD-C/ADHD-I subjects whose data is archival (and was, therefore, administered by others not associated with this study). Given the lack of information regarding how the subjects were administered the TOVA by others, it is difficult to make any particular conclusion in this regard. However, it remains one possible source of error.

It is difficult to determine the role of each of these factors in contributing to the results of this study. It does appear to be the case that this study had insufficient power to uncover some true effects (although the sample sizes in this study are similar to that of previous research in this area – e.g., Barkley & Grodzinsky, 1994). Many of this study's other limitations result from its archival nature. For example, the manner in which subjects were administered the TOVA could have been more closely monitored in a non-archival experimental design.

Of course, the strength of archival data is that it may more accurately reflect the type of clients who present themselves for evaluation in the "real world" (in contrast to those who choose to present themselves for a research project). To better understand the "real world" usefulness of the TOVA in differentially diagnosing ADHD subtypes from LD children and non-patient controls, the sensitivity, PPP, and NPP of the major variables was computed. In order to be useful in differentially diagnosing a condition such as ADHD, a measure should be able to demonstrate (at the very least) better than chance classification of a subjects into their respective groups. As applied to PPP/NPP, results should therefore be considered useful only if they exceed .50 (chance performance), with greater diagnostic confidence being given to those variables whose PPP/NPP values are closer to 1.0 (i.e., perfect classification). Of course, the PPP/NPP values for a measure should be interpreted in the context of its sensitivity. It is possible, for instance, that a variable could have nearly perfect classification ability (i.e., PPP values approaching 1.0), but not detect many cases of a disorder.

The results of this study suggested poor sensitivity for commission errors and omission errors, moderate sensitivity for response time, and good sensitivity for response time variability (with 78% of ADHD children displaying an abnormal score on this measure). Indeed, the response time variability measure appeared to be easily the most useful score, particularly when the ADHD groups were combined and the LD group was eliminated from the classification scheme. In this situation, the PPP was 88%, and the NPP 60%. Such results appear to support the use of this measure in the clinical diagnosis of ADHD, particularly when the presence of a learning disorder has been ruled-out through other testing. Abnormal findings seem to suggest the presence of ADHD (though not of which subtype).

Another promising measure from the TOVA is response time. The sensitivity of this variable is acceptable (61% of those diagnosed with ADHD had abnormal findings on this measure), and it's PPP was 83% (i.e., 83% of positive findings involved children diagnosed with ADHD) when the ADHD groups were combined and LD was eliminated from the classification scheme. The NPP was a far weaker 33% in these circumstances, suggesting that a negative finding should not be interpreted. When the LD group was included in the classification scheme, the PPP for the combined ADHD groups falls to 67% (with a 59% NPP). In any case, a positive finding on this measure appears to accurately detect the presence of ADHD (though not of which subtype), particularly once disorders such as LD have been ruled-out by separate means.

The use of commission errors appears somewhat more problematic. The sensitivity of this measure was a mere 22% (suggesting that only 22% of those diagnosed with ADHD had an abnormal finding on this measure). Its PPP was 83% when combining the ADHD groups and removing the LD group from consideration. However, the NPP was only 33% in these circumstances, suggesting that normal findings on this measure are of little interpretive value.

Finally, the omission errors measure had a 39% sensitivity to ADHD. This variable had an 82% PPP when the ADHD groups were combined and LD removed from consideration (82% of abnormal findings were from ADHD subjects). Thus, abnormal findings on this measure tend to suggest the presence of ADHD, once disorders such as a learning disability have been ruled-out. However, the NPP on this measure (36% under these circumstances) suggests that normal results can generally not be trusted under these conditions.

The second hypothesis in this study was that the TOVA would be useful in differentially diagnosing ADHD-I from ADHD-C; in particular, it was thought that ADHD-C children would commit more errors of commission, while ADHD-I children would commit more errors of omission, have slower response times, and display greater response time variability. Results from this study failed altogether to support these conclusions, a finding which is consistent with a previous failure to do so with an earlier version of the CPT and other measures (Barkley & Grodzinsky, 1994). Given the evidence suggesting differing patterns of behavioral disturbance and contrasting family psychiatric history, it would be premature to conclude that these two disorders are, indeed, two subtypes of a similar disorder. Nevertheless, it is clear that ADHD-C and ADHD-I children perform similarly on Continuous Performance Tests like the TOVA. It may well be that our limited understanding of attention (and correspondent limited ability to measure its different facets) is impeding our ability to differentially diagnose these conditions with tests such as the TOVA. This may be considered quite ironic, given the fact that attention has been researched for over 100 years (Barkley, 1990). However, our increased sophistication in understanding attention (as exemplified in the development of "types" of attention – such as "sustained attention" or "selective attention") has led to yet further questions about the interconnectedness between "attention" and other brain functions, as well as regarding the many brain dysfunctions which can affect attentional functioning. Perhaps our ability to differentially diagnose varying types of attentional disorders will remain limited until we more fully understand these subtle and complex interactions.

The final hypothesis investigated in this study was that a "High Consistency" subgroup of ADHD children would perform better than controls on all TOVA variables. This hypothesis was generated by examining the work with elderly subjects by Ferraro and Moody (1996) which indicated that, while some elderly do experience a decline in mental processing speed, a subset of elderly with high-consistency response times (i.e., low standard deviations in their response times) actually outperformed younger subjects on a measure of choice reaction time. By analogy, it was though that a subset of high-consistency ADHD children (i.e., those with low response time variability) might outperform controls on the other TOVA variables. In favor of this hypothesis were the following findings: (1) "High Consistency" children committed fewer omission errors than "Low Consistency" children, (2) "High Consistency" children committed fewer commission errors than "Low Consistency" children; and (3) "Low Consistency" children had slower response times than "High Consistency" children. However, no interactions which would have supported this hypothesis (e.g., an interaction between group and consistency suggesting that "High Consistency" ADHD children outperformed other groups) were present on any of the TOVA variables.

Overall, then, this study provided little support for the hypothesis that "High Consistency" ADHD children would outperform "Low Consistency" ADHD children and controls. This may be due in part to the fact that fewer of the ADHD group members fell in the "High Consistency" category, and those who did tended to fall closer to the median. Given the finding that consistency overall significantly affected scores on major TOVA variables, it should perhaps not be surprising that this hypothesis was generally unsupported.

Finally, it is interesting to note the lack significant interactions in this study, despite their presence in a previous study using the TOVA by Clay et al. (1996). For instance, the Clay study reported a significant group by quartile interaction on the omission errors variable – with increasing discrepancies between ADHD and non-ADHD subjects on this variable in the third and fourth quartiles. The present study failed to repeat this finding. Further, the Clay study reported significantly decreasing discrepancies between LD and non-LD subjects on the response-time variable across quartiles. Again, this pattern of findings was not repeated in this study. It should be noted, of course, that the Clay study (which compared ADHD, ADHD+LD, LD, and non-patient controls) involved comparisons between groups which were not studied in the present investigation (e.g., children with comorbid ADHD and RD were not included in the present investigation).

Limitations

Several limitations of this study should be kept in mind during its interpretation. First of all, the sample sizes were somewhat small, with 15 ADHD-I subjects, 20 LD subjects, 20 non-patient controls, and 31 ADHD-C children. Although of similar sample size to other published projects in this area (e.g., Barkley & Grodzinsky, 1994), this study was shown to have insufficient power to detect some true effects (of one standard deviation or above). Second, this study is limited by its primarily archival nature. Various clinicians used their separate judgments in determining which children met the DSM-IV diagnostic criteria for ADHD, and standardized assessment measures for ADHD were not used in this study because they were not always used by the diagnosing clinician. An attempt was made to compensate for this somewhat by including in my study only those children whose clinical chart contained sufficient documentation of enough DSM-IV criteria to merit a diagnosis of ADHD. Nevertheless, the use of a variety of clinicians and a reliance on clinician judgment may have produced somewhat impure groups (i.e., groups within which the severity of attentional difficulties may have been quite variant). Finally, in another disadvantage of archival research, I was unable to observe the archival data children (i.e., those in the ADHD-C and ADHD-I groups) during their TOVA administration (to ensure that the test instructions and test environment were similar – as well as to observe their behavior during the test). Certainly, TOVA results which were assessed to be blatantly invalid by the TOVA interpretation program (the TOVA test interpretation program does this when a sufficient number of obvious omission errors occur, for example) or the clinician who observed the child (e.g., if the clinician noted that the child refused to follow test instructions during the last five minutes of the TOVA) were not included in this study. Nevertheless, the lack of standardization and opportunity to observe the children during their assessment may have contributed to impure subgroups.

Future research with the TOVA may wish to focus on addressing these concerns. In other words, it would appear logical to conduct a study with the TOVA in one setting, using objective measures of attention deficits, and to use a larger sample size. It would also be interesting to include a group of ADHD, primarily hyperactive/impulsive subtype children. Such a study would likely be useful in better distinguishing the pattern of CPT performance generated by inattention in contrast to hyperactivity/impulsivity. Finally, given the family psychiatric history of depression commonly found in those diagnosed with ADHD-I, it would be useful to examine the TOVA performance of depressed children. This might help to either support the similarity of ADHD-I to childhood depression – or serve as a useful tool in differentiating these conditions.

REFERENCES

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Aman, M. G., & Turbott, S. H. (1986). Incidental learning, distraction, and sustained attention in hyperactive and control subjects. *Journal of Abnormal Child Psychology, 14*, 441-445.

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

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd edition, revised).* Washington, DC: Author.

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

Barkley, R. A. (1990). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York, NY: Guilford Press.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1990). Comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. *Journal of Consulting and Clinical Psychology, 58(6)*, 775-789.

Barkley, R. A., DuPaul, G. J., & McMurray, M. B. (1991). Attention deficit disorder with and without hyperactivity: Clinical response to three dose levels of methylphenidate. *Pediatrics, 887*, 519-531.

Barkley, R. A., & Grodzinsky, G. M. (1994). Are tests of frontal lobe functions useful in the diagnosis of attention deficit disorders? *The Clinical Neuropsychologist, 8(2)*, 121-139.

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.

Bender, L. (1938). A Visual-Motor Gestalt Test and its clinical use. *American Orthopsychiatric Association*. (Research Monograph No. 3).

Benton, A. L., & Hamsher, D. deS. (1978). *Multilingual aphasia examination* (manual: revised). Iowa City: University of Iowa.

Berry, C. A., Shaywitz, S. E., & Shaywitz, B. A. (1985). Girls with attention deficit disorder: A silent minority? A report on behavioral and cognitive characteristics. *Pediatrics, 76*, 801-809.

Borys, S. V., Spitz, H. H., & Dorans, B. A. (1982). Tower of Hanoi performance of retarded and non-retarded students. *American Journal of Mental Deficiency, 84*, 280-288.

Boucagnani, L. L., & Jones, R. W. (1989). Behaviors analogous to frontal lobe dysfunction in children with attention deficit hyperactivity disorder. *Archives of Clinical Neuropsychology, 4*, 161-173.

Campbell, J. W., D'Amato, R. C., Raggio, D. J., & Stephens, K. D. (1991). Construct validity of the computerized Continuous Performance Test with measures of intelligence, achievement, and behavior. *Journal of School Psychology*, *29*, 127-134.

Campbell, S.B. & Werry, J. S. (1986). Attention deficit disorder (hyperactivity). In H.C. Quay and J. S. Werry (Eds.), *Psychopathological Disorders of Childhood* (3rd ed.) New York: Wiley.

Carlson, C. L., Lahey, B. B., & Neeper, R. (1986). Direct assessment of the cognitive correlates of attention deficit disorders with and without hyperactivity. *Journal of Behavioral Assessment and Psychopathology*, *8*, 69-86.

Cerella, J. (1994). Generalized slowing in Brinley plots. *Journal of Gerontology: Psychological Sciences, 49*, 65-71.

Chelune, G. J., Ferguson, W., Koon, R., & Dickey, T. O. (1986). Frontal lobe disinhibition in attention deficit disorder. *Child Psychiatry and Human Development*, *16*, 221-234.

Clay, D., Petros, T., Searcy, E., & Westby, S. (1996). TOVA utility in the differential diagnosis of attention deficit hyperactivity disorder and learning disability. Manuscript in preparation.

Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, *126*, 884-888.

Conners, C. K. (1979a). *Conners Parent Rating Scale* (Hyperkinesis Index). Chicago, IL: Abbott Laboratories.

Conners, C. K. (1979b). *Conners Teacher Rating Scale* (Hyperkinesis Index). Chicago, IL: Abbott Laboratories.

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

Dainer, K. B., Klorman, R., Salzman, L. F., Hess, D. W., Davidson, P. W., & Michael, R. L. (1981). Learning-disordered children's evoked potentials during sustained attention. *Journal of Abnormal Child Psychology, 9*, 79-84.

Denckla, M. B. (1991). Attention deficit hyperactivity disorder: Residual type. *Journal of Child Neurology,* 6(Suppl.), S44-S50.

Douglas, V. I. (1983). Attentional and cognitive problems. In M. Rutter (Ed.), *Developmental neuropsychiatry*. New York: Guilford.

Dunn, L. M., & Dunn, L. (1981). *Peabody Picture Vocabulary Test – Revised.* Circle Pines, MN: American Guidance Service.

Edelbrock, C., Costello, A., & Kessler, M. D. (1984). Empirical corroboration of attention deficit disorder. *Journal of the American Academy of Child Psychiatry, 23*, 285-290.

Ferraro, F. R., Moody, J. (1996). Consistent and inconsistent performance in young and elderly adults. *Developmental Neuropsychology, 12(4)*, 429-441.

Fisk, A. D., & Fisher, D. L. (1994). Brinley plots and theories of aging: The explicit, muddled, and implicit debates. *Journal of Gerontology: Psychological Sciences, 49*, 81-89.

Gordon, M. (1983). *The Gordon Diagnostic System*. DeWitt, NY: Gordon Systems.

Goyette, C. H., Conners, C. K. & Ulrich, R. F. (1978). Normative data on revised Conners parent and teacher rating scales. *Journal of Abnormal Child Psychology, 6*, 221-236

Greenberg, L. M., & Waldman, I. D. (1993). Developmental normative data on the Test of Variables of Attention (TOVA). *Journal of Child Psychology and Psychiatry and Allied Disciplines, 34(6)*, 1019-1030.

Halperin, J. M., Newcorn, J. H., Sharma, V., Healey, J. M., Wolf, L. E., Pascualvaca, D. M., & Schwartz, S. (1990). Inattentive and noninattentive ADHD children: Do they constitute a unitary group? *Journal of Child Abnormal Psychology, 18(4)*, 437-449.

Halperin, J. M., Sharma, V., Greenblatt, E., & Schwartz, S. (1991). Assessment of the continuous Performance Test: Reliability and validity in a nonreferred sample. *Journal of Consulting and Clinical Psychology, 3*, 603-608.

Halperin, J. M., Wolf, L. E., Pascualvaca, D. M., Newcorn, J. H., Healey, J. M., O'Brien, J. D., Morganstein, A., & Young, G. (1988). Differential assessment of attention and impulsivity in children. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 326-329.

Heaton, R. K. (1981). *A manual for the Wisconsin Card Sorting Test*. Odessa, FL: Psychological Assessment Resources.

Hechtman, L. (1991). Developmental, neurobiological, and psychosocial aspects of hyperactivity, impulsivity, and inattention. In M. Lewis (Ed.), *Child and adolescent psychiatry. A comprehensive textbook.* Baltimore, MD: Williams & Wilkins.

Hynd, G. W., Lorys-Vernon, A. R., Semrud-Clikeman, M., Nieves, N., Huettner, M. I., & Lahey, B. B. (1993). Attention deficit disorder without hyperactivity (ADD/WO): A distinct behavioral and neurocognitive syndrome. *Journal of Child Neurology, 6*, 35-41.

Jastak, S., & Wilkinson, G. S. (1984). *Wide Range Achievement Test – Revised.* Wilmington, DE: Jastak Associates.

Kagan, J. (1964). *The Matching Familiar Figures Test.* Cambridge: Harvard University.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test.* Malvern, PA: Lea & Febiger.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman assessment battery for children – Interpretive manual.* Circle Pines, MN: American Guidance Service.

King, C., & Young, R. (1982). Attentional deficits with and without hyperactivity: teacher and peer perceptions. *Journal of Abnormal Child Psychology, 10*, 483-496.

Kirk, R. E. (1982). *Experimental Design: Procedures for the Behavioral Sciences, 2nd Edition.* Monterey, CA: Brooks-Cole.

Klee, S. H., & Garfinkel, B. D. (1983). The computerized continuous performance test: a new measure of inattention. *Journal of Abnormal Child Psychology, 11(4)*, 487-496.

Kraemer, H. C., & Theimann, S. (1987). *How Many Subjects?* Newbury Park, CA: Sage.

Kupietz, S. S. (1990). Sustained attention in normal and in reading-disabled youngsters with and without ADDH. *Journal of Abnormal Child Psychology, 18(4)*, 357-372.

Lahey, B. B. (October, 1988). *Attention deficit disorder without hyperactivity: Issues of validity.* Paper presented to the annual meeting of the Bloomingdale Conference on Attention Deficit Disorder, Seattle, WA.

Lahey, B. B., Pelham, W. E., Schaughency, E. A., Atkins, M. S., Murphy, H. A., Hynd, G. W., Russo, M, Hartdagen, S., & Lorys-Vernon, A. (1988). Dimensions and types of attention deficit disorder with hyperactivity in children: A factor and cluster analytic approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 330-335.

Lahey, B. B., Shaughency, E., Hynd, G., Carlson, C., & Nieves, N. (1987). Attention deficit disorder with and without hyperactivity: Comparison of behavioral characteristics of clinic-referred children. *Journal of the American Academy of Child Psychiatry, 26*, 718-723.

Lahey, B. B., Shaughency, E., Strauss, C., & Frame, C. (1984). Are attention deficit disorders with and without hyperactivity similar or dissimilar disorders? *Journal of the American Academy of Child Psychiatry, 23*, 302-309.

Lassiter, K. S., D'Amato, R. C., Raggio, D. J., Whitten, J. C. M., & Bardos, A. N. (1994). The construct specificity of the continuous performance test: does inattention relate to behavior and achievement? *Developmental Neuropsychology, 10(2)*, 179-188.

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

Lindgren, S. D., & Lyons, D. A. (1984). *Pediatric Assessment of Cognitive Efficiency (PACE)*. Iowa City: University of Iowa, Department of Pediatrics.

Loge, D. V., Staton, R. D., & Beatty, W. W. (1990). Performance of children with ADHD on tests sensitive to frontal lobe dysfunction. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 540-545.

Lovejoy, M. C., & Rasmussen, N. H. (1990). The validity of vigilance tasks in differential diagnosis of children referred for attention and learning problems. *Journal of Abnormal Child Psychology, 18(6)*, 671-681.

McIntyre, C. W., Blackwell, S. L., & Denton, C. L. (1978). Effect of noise distractibility on the spans of apprehension of hyperactive boys. *Journal of Abnormal Child Psychology, 6*, 483-491.

Margolis, J. S. (1972). *Academic correlates of sustained attention*. Unpublished doctoral dissertation. University of California at Los Angeles.

Markwardt, F. C. (1989). *Peabody Individual Achievement Test – Revised.* Circle Pines, MN: American Guidance Service.

Matier-Sharma, K., Perachio, N., Newcorn, J. H., Sharma, V., & Halperin, J. M. (1995). Differential diagnosis of ADHD: Are objective measures of attention, impulsivity, and activity level helpful? *Child Neuropsychology, 1(2)*, 118-127.

Maurer, R. G., & Stewart, M. (1980). Attention deficit disorder without hyperactivity in a child psychiatry clinic. *Journal of clinical psychiatry, 41*, 232-233.

Nărhi, V., & Ahonen, T. (1995). Reading disability with or without attention deficit hyperactivity disorder: do attentional problems make a difference? *Developmental Neuropsychology, 11(3)*, 337-349.

Nuechterlein, K. H. (1983). Signal detection in vigilance tasks and behavioral attributes among offspring of schizophrenic mothers and among hyperactive children. *Journal of Abnormal Psychology, 92*, 4-28.

O'Dougherty, M, Nuechterlein, K. H., & Drew, B. (1984). Hyperactive and hypoxic children: Signal detection, sustained attention, and behavior. *Journal of Abnormal Psychology, 93*, 178-191.

Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology, 29(3)*, 511-523.

Porteus, S. D. (1965). *Porteus Maze Test: Fifty years application*. New York: Psychological Corporation.

Radosh, A., & Gittelman, R. (1981). The effect of appealing distracters on the performance of hyperactive children. *Journal of Abnormal Child Psychology, 9*, 179-189.

Raggio, D. (1991). *Raggio Evaluation of Attention Deficit Disorder* [Computerized Test]. Jackson: University of Mississippi Medical Center, Infant and Child Development Clinic.

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

Rosenthal, R. H., & Allen, T. W. (1978). An examination of attention, arousal, and learning dysfunctions of hyperkinetic children. *Psychological Bulletin, 85*, 689-715.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, Jr. & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*, 343-350.

Rubinstein, R. A., & Brown, R. T. (1984). An evaluation of the validity of the diagnostic category of attention deficit disorder. *American Journal of Orthopsychiatry, 54*, 398-414.

Schaughency, E. A., Lahey, B. B., Hynd, G. W., Stone, P. A., Piacentini, J. C., & Frick, P. J. (1989). Neuropsychological test performance and the attention deficit disorders: Clinical utility of the Luria-Nebraska Neuropsychological Battery – Children's Revision. *Journal of Consulting and Clinical Psychology, 57*, 112-116.

Sergeant, J. A., & Scholten, C. A. (1985). On resource strategy limitations in hyperactivity: Cognitive impulsivity reconsidered. *Journal of Child Psychology and Psychiatry, 26*, 97-109.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Taylor, E. (1994). Syndromes of attention deficit and hyperactivity. In M. Rutter, E. Taylor, & L. Hersov (Eds.), *Child and adolescent psychiatry: Modern approaches*. New York: Blackwell Scientific.

Taylor, H. H., Fletcher, J. M., & Staz, P. (1984). Neuropsychological assessment of children. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (pp. 211-234). New York: Pergamon Press.

Ullman, R. K., Sleator, E. K., & Sprague, R. L. (1986). *ADD-H Comprehensive Teacher's Rating Scale*. Chicago, IL: MetriTech, Inc.

van der Meere, J., Wekking, E., & Sergeant, J. (1991). Sustained attention and pervasive hyperactivity. *Journal of Child Psychology and Psychiatry, 32*, 275-284.

Wechsler, D. (1974). *Wechsler intelligence scale for children – revised.* New York: Psychological Corporation.

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

Weyandt, L. L., & Willis, W. G. (1994). Executive functions in school-aged children: potential efficacy of tasks in discriminating clinical groups. *Developmental Neuropsychology, 10(1)*, 27-38.

Whalen, C. K. (1989). Attention deficit and hyperactivity disorders. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child psychopathology*. New York: Plenum.

White, D. M., & Sprague, R. L. (1992). The "attention deficit" in children with attention-deficit hyperactivity disorder. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology, Vol. 14*. New York: Plenum.

Wicks-Nelson, R., & Israel, A. (1997). *Behavior disorders of childhood (3rd edition)*. Upper Saddle River, NJ: Prentice Hall.

Woodcock, R. W., & Johnson, M. B. (1989, 1990). *Woodcock-Johnson Psychoeducational Battery – revised*. Denton, TX: DLM Resources.

Zametkin, A. J., & Rapoport, J. L. (1986). The pathophysiology of attention deficit disorder with hyperactivity: a review. In B. B. Lahey & A. E. Kazdin (Eds.) *Advance in clinical child psychology, Vol. 9*. New York: Plenum.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989