Child Neuropsychology
2000, Vol. 6, No. 2, pp. 101-114
0929-7049/00/0602-101$15.00

The Predictive Power of Combined Neuropsychological Measures for Attention-Deficit/Hyperactivity Disorder in Children

Eve Marie Perugini[1], Elizabeth A. Harvey[2], David W. Lovejoy[3], Kristen Sandstrom[1], and Anne H. Webb[1]
[1]University of Connecticut, [2]University of Massachusetts, and [3]Psychology Testing Department, Hartford Hospital and The Institute of Living, Connecticut

ABSTRACT

The present study explores the predictive power of seven neuropsychological assessment tools used in combination in classifying children with attention-deficit/hyperactivity disorder (ADHD). Twenty-one ADHD boys and 22 community control children participated. Group differences were significant on the continuous performance test only; however, battery analysis did increase overall predictive power, which was moderate. This study highlights the difficulty in identifying consistent mean differences on tests of frontal/executive functioning across studies, as well as the need to assess the predictive validity of these tests in classifying children with ADHD. The study suggests that these tests may provide greater predictive validity when used in combination. Inconsistencies in the literature are discussed, with consideration of research methodology, the heterogeneity of the ADHD population, and comorbid diagnoses.

Attention-deficit/hyperactivity disorder (ADHD) is a developmental disorder which is characterized by inattention, hyperactivity, and impulsivity. ADHD occurs in approximately five percent of the childhood population, with a higher incidence in boys than girls by a ratio of approximately three to one (Barkley, 1998). The current method of diagnosis involves extensive parent interview and parent- and teacher-completed behavior rating scales. Researchers have long sought more objective measures that might be individually administered to determine the presence of ADHD, with limited success. The National Institutes of Health (NIH) recently convened a consensus panel on the diagnosis and treatment of ADHD which highlighted inconsistency in ADHD diagnosis and the need for further validation of the disorder (National Institutes of Health, 1998).

Research suggests that ADHD may have a neurophysiological basis, involving frontal cortical regions and other interconnected subcortical structures (Boucagnani & Jones, 1989; Castellanos, 1997). These findings are consistent with models of attention which emphasize the frontal lobes (Luria, 1966; Posner & Dehaene, 1994). In addition, a convergence of evidence from studies of neuroanatomy, neuroimagery, neurochemistry, and stimulant medication has led many researchers to view ADHD as a frontal lobe dysfunction (Barkley, 1997; Castellanos, 1997; Evans, 1987).

The tests believed to assess frontal functioning include measures of the ‘‘executive functions,’’ such as selective and sustained attention, inhibition of responses, working memory, organization, and planning of complex behaviors (Lezak, 1995). Evaluation of the performance of

Address correspondence to: Elizabeth Harvey, Department of Psychology, University of Massachusetts, Tobin Hall, P.O. Box 37710, Amherst, MA 01003-7710. E-mail: eharvey@psych.umass.edu
Accepted for publication: January 5, 2000.

ADHD children on frontal lobe/executive tasks has produced mixed results. Barkley, Grodzinsky, and DuPaul (1992) reviewed 22 neuropsychological studies of frontal functions in ADHD children and reported that impairment tended to be inconsistent across studies. Their review revealed a limited number of frontal/executive tests that reliably differentiated ADHD children from controls, and a recent study by two of these authors (Grodzinsky & Barkley, 1999) has suggested greatest reliability for the Continuous Performance Tasks (CPTs; e.g., Conners, 1995), the Stroop Color-Word Association Test (Stroop, 1935), the Controlled Oral Word Association Test (COWAT, or F-A-S; Benton & Hamsher, 1978), and the Hand Movements Scale from the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). The authors suggest that tests of response inhibition distinguish ADHD children with hyperactive symptoms more reliably than other tests of frontal lobe dysfunctions (Barkley et al., 1992).

Three of these measures have quite consistently differentiated ADHD from non-ADHD children: the CPTs (e.g., Fischer, Barkley, Edelbrock, & Smallish, 1990; Grodzinsky & Barkley, 1999; Grodzinsky & Diamond, 1992; Mariani & Barkley, 1997), the Stroop (e.g., Boucagnani & Jones, 1989; Carlson, Lahey, & Neeper, 1986; Carter, Krener, Chaderjian, Northcutt, & Wolfe, 1995; Gorenstein, Mammato, & Sandy, 1989; Grodzinsky & Barkley, 1999; Grodzinsky & Diamond, 1992; Hopkins, Perlman, Hechtman, & Weiss, 1979; Seidman, Biederman, Faraone, Weber, & Ouellette, 1997), and the Hand Movements Scale (e.g., Breen, 1989; Grodzinsky & Diamond, 1992; Mariani & Barkley, 1997). Many studies have also found an association between ADHD and the COWAT (e.g, Felton, Wood, Brown, & Campbell, 1987; Grodzinsky & Barkley, 1999; Grodzinsky & Diamond, 1992), the Trail Making Test (e.g, Boucugnani & Jones, 1989; Chelune, Ferguson, Koon, & Dickey, 1986; Gorenstein et al., 1989; Johnson, 1991; Moffitt & Silva, 1988; Shue & Douglas, 1992), and the Freedom From Distractibility (FFD) factor of the Wechsler Intelligence Scale for Children (e.g., Anastopoulos, Spisto, and Maher, 1994; Loge, Staton, & Beatty, 1990; Cohen, Becker, & Campbell, 1990). However, there have also been many studies that have failed to support the COWAT, Trails, and the FFD factor (e.g., Fischer et al., 1990; Grodzinsky & Barkley, 1999; Grodzinsky & Diamond, 1992; Loge et al., 1990; Mariani & Barkley, 1997; McGee, Williams, Moffitt, & Anderson, 1989; Reader, Harris, Schuerholz, & Denckla, 1994; Riccio, Cohen, Hall, and Ross, 1997; Semrud-Clikeman, Hynd, Lorys, and Lahey 1994).

While several studies revealed significant mean differences between ADHD and control groups on neuropsychological measures, they often failed to examine the accuracy with which these tests could diagnose ADHD. Classification analyses of the tests, including sensitivity, specificity, positive predictive power (PPP) and negative predictive power (NPP) must be evaluated in order to determine each test's ability to classify children as having ADHD. Sensitivity describes the proportion of individuals who have a certain disorder whose score on the measure suggests impairment. Specificity refers to the proportion of non-disordered individuals who do not have an impaired score on the instrument. Positive predictive power describes the proportion of individuals who receive an impaired score who have the disorder. Negative predictive power refers to the proportion of individuals who do not have an impaired score who do not have the disorder. Overall predictive power (OPP) refers to the proportion of individuals who are correctly classified by a given test. Finally, an odds-ratio refers to the ratio of the odds of a disordered individual receiving an impaired score to the odds of non-disordered individuals receiving an impaired score. The odds-ratio provides an index that is not influenced by the base rate of the disorder.

Two recent studies have examined more closely the accuracy of neuropsychological measures in diagnosing ADHD. Grodzinsky and Barkley (1999) attempted to classify children with ADHD by using a battery of tests assessing frontal lobe functions. A subset of the tests, including the CPT, COWAT, Hand Movements Scale, and the Stroop Color-Word Association Test each had good positive predictive power in identifying children with ADHD, but poor negative predictive power; an abnormal score was predictive of the presence of ADHD, but a normal score could not reliably rule out the disorder. Similarly, Doyle, Biederman, Seidman, Weber, and Faraone (2000) found good positive predictive power and specificity, but poor sensitivity and negative predictive power for the Stroop Color-Word Association Test, Wisconsin Card Sort Test, Wide Range Assessment of Memory and Learning, California Verbal Learning Test, Auditory CPT, Letter Cancellation, Rey Osterreith Complex Figure, and the Wechsler Freedom from Distractibility Index. Both studies concluded that the tests had some clinical value, but that their limited negative predictive power, poor sensitivity, and an overall fair-to-poor classification of children having ADHD made them ineffective tools for diagnostic use.

Despite these limited findings, other researchers continue to seek a neuropsychological measure of ADHD. Lovejoy et al. (1999) compared a group of ADHD adults to normal controls on six neuropsychological measures believed to be sensitive to frontal lobe functioning. The tests included the COWAT, the California Verbal Learning Test, the Stroop Neuropsychological Screening Test, the Trail Making Test Part A & B (Reitan & Wolfson, 1985), and the Wechsler Adult Intelligence Scale-Revised Freedom from Distractibility Factor, comprised of the Digit Span and Arithmetic subtests. The majority of the measures discriminated between the two groups, but again these tests were compromised in terms of their negative predictive power. Lovejoy et al. then examined several ways of combining the results of these tests and identified a unique battery approach that yielded acceptable levels of sensitivity, specificity, positive predictive power, and negative predictive power. Specifically, they found that the following two criteria yielded overall classification rates of .90 and .83 respectively: (1) classifying an individual as ADHD if there was impairment on at least one out of the six tests and (2) classifying an individual as ADHD if there was impairment on at least two out of the six tests. Using stricter criteria requiring impairment on more than two tests decreased overall classification rates.

If a test battery approach can effectively discriminate between adults with and without ADHD, it might allow for easier, more economical, and more objective diagnostic classification of children as well. Doyle et al. (2000) evaluated this procedure with children and found some improvement but still limited predictive power. The present study explores the predictive power of several neuropsychological assessment tools used in combination in classifying children with ADHD. Measures were selected on the basis of their predictive power in previous studies of children (Grodzinsky & Barkley, 1999) and adults (Lovejoy et al., 1999). These tests of frontal/executive functioning include the K-ABC Hand Movements Scale, the Stroop Color-Word Association Test, the COWAT, the Trail Making Test Parts A and B, the Arithmetic and Digit Span subtests of the WISC-III (Wechsler, 1991), and Conners' CPT (Conners, 1995). It was hypothesized that children with ADHD would display significant impairment on many of these measures relative to controls, but as suggested by Lovejoy et al., predictive power would be greatest for the battery of tests considered collectively.

METHOD

Participants

The present study involved two groups: 21 boys previously diagnosed with ADHD, and 22 boys in a community control group. All children had participated in earlier studies conducted at the University of Connecticut and were recruited by phone. Previously administered measures included the Peabody Picture Vocabulary Test – Revised (PPVT-R; Dunn & Dunn, 1981), the Diagnostic Interview Schedule for Children – IV (DISC-IV; Shaffer, Fisher, Lucas, Dulcan, Schwab-Stone, 2000) and the Behavioral Assessment System for Children – Parent Report Scale (BASC-PRS; Reynolds & Kamphaus, 1992). The PPVT-R is a nonverbal, multiple-choice test designed to evaluate the receptive vocabulary of children and adults, and has been shown in numerous studies to have good reliability and validity as well as concurrent validity as a measure of cognitive ability (Dunn & Dunn, 1981; Breen, 1981). The DISC-IV is a structured interview based on the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 1994) that has demonstrated adequate validity and reliability of childhood diagnoses (Shaffer et al., 2000). The DISC-IV sections which diagnose ADHD, Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD) were used in the present study. The BASC-PRS is a comprehensive rating scale that assesses a broad range of psychopathology in children aged 2 years 6 months and older. This study examined the Hyperactivity and Conduct Problems subscales as completed by mothers. Research on the BASC suggests that it has excellent reliability and validity (Reynolds & Kamphaus, 1992), particularly for ADHD children (Ostrander, Wienfurt, Yarnold, & August, 1998).

Children met the following criteria to participate in the present study: All children had to (1) be between the ages of 6 and 12; (2) have an IQ estimate of 80 or higher based on the PPVT-R; and (3) have no evidence of deafness, blindness, color blindness, severe language delay, cerebral palsy, epilepsy, autism, learning disability, or psychosis as established through parent interview and child observations. Basic literacy was also required for participation, since one or more of the experimental tasks required the participant to read from a printed page.

Participants in the ADHD group were also required to: (1) meet DSM-IV criteria for ADHD Combined Type or Hyperactive/Impulsive Type based on the DISC-IV; (2) have a T-score greater than or equal to 65 on the BASC-PRS Hyperactivity Subscale (mothers' ratings); and (3) have no history of treatment with stimulant drugs or, if such a history, agree to be removed from medication for 24 hours prior to evaluation in the study. If parents had any concerns about discontinuing medication, they were invited to discuss this with their prescribing physician. This is a standard procedure in research on ADHD, and has been used in previous studies (e.g., Barkley & Grodzinsky, 1994). Testing was conducted on weekends or during the summer months so that discontinuing medication did not interfere with school activities. Participants taking psychotropic medications other than stimulants were excluded from participation.

Children were recruited from area child care centers in order to create a control group similar to the ADHD group in gender, age, and IQ. Participants in the control group had no history of mental health services for behavioral or emotional problems, no parent or teacher complaints of significant behavior problems, T-scores on the BASC-PRS Hyperactivity Subscale below 60, and no history of academic problems as reported by mothers. All study participants were male children from central Connecticut, with a mean age of 116 months (age 9 years, 8 months) in the ADHD group and 110 months (age 9 years, 2 months) in the control group. The mean IQ estimate based on the PPVT-R was 110 and 114 for the ADHD and control groups respectively. Of the ADHD participants, 16 out of 21 (76%) met criteria for a comorbidity diagnosis of either ODD or CD as indicated by the DISC-IV. This proportion is consistent with various studies showing that 54–67% of ADHD children meet diagnostic criteria for ODD, and 20–56% of ADHD children meet criteria for CD (Barkley, 1998).

Procedure

Parents completed a brief phone interview to determine eligibility for the study. Children selected for the study were then administered approximately one hour of neuropsychological tests in the home, and were paid $10 for their participation. Parents signed a consent form prior to testing. Tests were administered by the student project director or an undergraduate research assistant trained in the administration of the tests. After a brief introduction, participants were administered seven neuropsychological tests of frontal/executive functioning. These dependent measures were selected on the basis of their predictive power in previous studies of children (Grodzinsky & Barkley, 1999) and adults (Lovejoy et al., 1999). Tests of frontal/executive functioning were administered in the following order:

Kaufman Hand Movements Scale (from the K-ABC; Kaufman & Kaufman, 1983)

This test measures the participant's ability to imitate progressively longer sequences of skilled hand movements. Raw scores are based on the number of sequences performed correctly, and standard scores ($M = 10$, $SD = 3$) based on performance and age were used for analysis. Research has demonstrated satisfactory reliability and validity for the K-ABC (Kaufman & Kaufman, 1983).

The Stroop Color-Word Association Test (Stroop, 1935)

This test includes three tasks, each corresponding to a separate card containing five columns of 20 items. In the first task, the participant must read a list of color names (red, blue, and green) written in black ink as quickly as possible. In the second task, the participant identifies colored patches of ink (also red, blue, and green) as quickly as possible. During the third, or "interference" task, the participant must identify the color of ink in which a word is printed as quickly as possible, the ink color being incongruent with the printed word. The participant is allowed 45 seconds for each of the three tasks, and scores are based on the number of items completed during each task.

The Controlled Oral Word Association Test (COWAT, or F-A-S; Benton & Hamsher, 1978)
Originally developed as part of Benton and Hamsher's Multilingual Aphasia Examination (MAE), this measure of verbal fluency requires the participant to make oral associations to three different letters of the alphabet, naming words beginning with "F", then "A" and "S," excluding plurals. One minute is allowed for each stimulus letter, and the total score is the sum of all acceptable words produced within the time limit for all three letters.

Trail Making Test (Parts A and B; Reitan & Wolfson, 1985)
This test is administered in two parts (A and B). A version designed for children aged 8–12 was used for the study. In Part A, the participant must use a pencil to connect a series of numbered circles distributed arbitrarily on a printed page. In Part B, the printed circles contain either letters or numbers, and the participant must alternate between numbers and letters in proper sequence (e.g., 1-A-2-B, and so on). Scores are based on the time to complete either Part A or Part B.

Arithmetic and Digit Span subtests of the Wechsler Intelligence Scale for Children, Third Edition (WISC-III, Freedom from Distractibility Factor; Wechsler, 1991)
These two subtests from the WISC-III are believed to assess attention skills and working memory. The Arithmetic subtest measures basic numeric skill, and is comprised primarily of oral word problems along with some easier counting items and more difficult written word problems. The Digit Span subtest is in two parts; Digits Forwards measures the participant's ability to repeat progressively longer sequences of orally presented numbers. Digits Backwards requires the participant to repeat progressively longer sequences of orally presented numbers in reverse sequence. The raw score is based on the total number of correct sequences on both tasks. Standard scores ($M = 10$, $SD = 3$) based on performance and participant age were used for further analysis of both the Arithmetic and Digit Span subtests. Numerous studies have shown good reliability and validity for the WISC-III (Wechsler, 1991).

Conners' Continuous Performance Test, (or CPT; Conners, 1995)
This is a 14-minute vigilance test during which letters are presented on a computer display screen. Children must press the space bar following each letter, but refrain from pressing the space bar when the letter "X" appears on the screen. Performance variables include the number of correct responses ("hits"), omissions ("misses"), and commissions (pressing the space bar inappropriately). An overall index score was used for analysis; this score is a composite index of 11 performance measures including hit reaction time, omission errors, commission errors, and several measures of variability of performance. The test developers report good sensitivity (.87) and specificity (.86) of the overall index, as well as satisfactory reliability and validity of this test (Conners, 1995).

RESULTS

An alpha of .01 was used in order to reduce Type 1 error. A series of univariate analyses of variance (ANOVA) suggested that the ADHD and non-ADHD groups did not differ significantly in age, F (1, 41) = .81, p = .37, or IQ as estimated by the PPVT-R, F (1, 41) = 1.22, p = .28. However, consistent with past research (Barkley, 1998), the non-ADHD group was significantly higher than the ADHD group in mother's education and family income and marginally higher in fathers' education (see Table 1). This difference may have been magnified in the present study by the method of recruitment for the control group. Control group children were recruited from community day care centers, and came primarily from two-income families.

Overall group differences on the seven neuropsychological measures were assessed through a multivariate analysis of variance (MANOVA). Differences between the ADHD and control groups were not significant on the seven measures when analyzed collectively, *Rao's R* (7, 34) = 1.83, p = .11; *Wilk's Lambda* = .73. Univariate analyses of variance (ANOVA) revealed significant group differences on one of

Table 1. Means and Standard Deviations by Group.

	ADHD (N = 21)		*Controls* (N = 22)			
	M	(*SD*)	*M*	(*SD*)	F (df)	
Age (in months)	115.67	(22.37)	110.14	(17.71)	.81	(1, 41)
PPVT	110.14	(12.20)	114.23	(12.07)	1.22	(1, 41)
Mother's education (years)	14.65	(2.08)	16.90	(2.49*)	10.07	(1, 40)*
Father's education (years)	14.24	(2.74)	16.64	(3.35a)	9.92	(1, 41)a
Family income (in thousands)	66.14	(30.05)	102.89	(42.54*)	9.92	(1, 39)*

* $p < .01$, [a] $p < .05$

the seven neuropsychological tests; ADHD children performed more poorly than their control counterparts on the CPT. Group differences on Trails B and Digit Span approached significance with ADHD children performing more poorly on these tasks. Group means and effect sizes are presented in Table 2.

Because of differences in performance demands between the Digits Forward and Digits Backwards tasks, these were also examined separately. Group differences approached significance on the Digits Forward task, F (1, 41) = 4.05, $p = .051$, but not on the Digits Backwards task, F (1, 41) = .03, $p = .87$. Although the CPT developers suggest the overall index as the most reliable indicator of test performance (Conners, 1995), previous studies have considered group differences in hits, omission errors, and commission errors, and these measures were also considered separately. Each approached significance (Hits: F (1,41) = 3.25, $p = .08$; Omissions: F (1, 41) = 3.07, $p = .09$; Commissions: F (1, 41) = 3.20, $p = .08$), but did not offer the level of prediction obtained from the overall index. ADHD children achieved fewer hits and made more omission errors than children in the control group, as predicted, but controls actually made more commission errors, which was not as predicted.

To identify impaired performance, a cutoff score on each of the seven tests was selected in order to analyze their predictive power in identifying children with ADHD. The overall index score of the CPT was considered impaired if it was greater than a value (nine) specified by test developers as indicating impaired performance (Conners, 1995). Norms for Trails B were not available, therefore a score 1.5 standard deviations below the mean of the control group was used for comparison, as has been used in previ-

Table 2. Univariate Analysis of Variance (ANOVA) Results of Differences Between ADHD and Control Groups on Neuropsychological Tests.

	ADHD (N = 21)		*Controls* (N = 22)			
	M	*SD*	*M*	*SD*	F (1, 41)	E.S.
Hand Movements (SS)	10.38	(2.25)	11.14	(2.59)	1.04	.31
Stroop Color-Word Trial (T)	49.38	(9.52)	52.59	(6.93)	1.61	.39
COWAT (FAS; normative)	5.48	(3.44)	6.23	(3.24)	.54	.22
Trails B (time)	63.29	(56.65)	42.18	(16.85)	2.80b	.51
Arithmetic (SS)	10.30	(2.89)	11.41	(3.02)	1.47	.38
Digit Span (SS)	9.00	(2.63)	10.73	(2.29)	5.29a	.70
CPT Overall Index	11.45	(7.30)	4.62	(5.29)	12.45*	1.07

Note. E.S. = effect size, SS = standard score
*$p < .01$, [a] $p < .05$, [b] $p = .10$

ous research with children aged six to twelve (Grodzinsky & Barkley, 1999). Although 1.5 standard deviations below the mean has been commonly used in previous research (e.g., Grodzinsky & Barkley, 1999), an examination of the distributions in the present study indicated that for the remaining tests, a cutoff of 1.5 standard deviations would have been too stringent because so few children in our sample performed below this cutoff. Furthermore, a recent study (Doyle, Biederman, Seidman, Weber, & Faraone, 2000) comparing the use of different cutoffs found that a 1 standard deviation cutoff produced better results than using more stringent cutoffs. Thus, for Hand Movements, the Arithmetic and Digit Span subtests of the WISC-III, and the Stroop Color-Word Trial, the cutoff score used was one standard deviation below the mean suggested by norms published by the test developers. For the COWAT (or FAS), a score was considered impaired if it fell below the 20th percentile based on published norms (Gaddes & Crockett, 1975).

Predictive power was then calculated for the battery of seven tests, as well as a battery including only the three tests which revealed significant or nearly significant differences between groups. Predictive power for each of these three tests (Digit Span, CPT, and Trails B) was also calculated separately. The results of these classification analyses are summarized in Table 3. For the battery of seven tests, an impaired score on at least two of seven neuropsychological tests provided the strongest prediction, Overall Predictive Power (OPP) = .77, with sensitivity, specificity, positive predictive power (PPP), and negative predictive power (NPP) ranging from .62 to .91. Furthermore, this approach produced an impressive odds-ratio of 16.25. Impaired scores on at least three of seven neuropsychological tests also provided significant overall predictive power, OPP = .65, but sensitivity was poor at .28.

The three-test battery of Digit Span, CPT, and Trails B produced similar results. An impaired score on at least one of three neuropsychological tests provided modest predictive power, OPP = .67, and was suggestive of ADHD (sensitivity = .76), but a finding of no impairment did not reliably rule out the disorder (specificity = .59). This approach provided a moderately high odds-ratio of 4.62. An impaired score on at least two of three neuropsychological tests provided greater overall prediction, OPP = .77, but sensitivity was weak (.52) given the more stringent classification criteria. Of the three tests consid-

Table 3. Predictive Power of a Seven-Test Battery, Three-Test Battery, and Three Individual Tests.

	Sensitivity	Specificity	PPP	NPP	OPP	ODDS	χ2
1+ out of 7	.76	.45	.57	.67	.60	2.67	2.22
2+ out of 7	.62	.91	.87	.71	.77	16.25	13.19**
3+ out of 7	.28	1.0	1.0	.59	.65	--b	7.31*
4+ out of 7	.05	1.0	1.0	.52	.53	--b	1.07
1+ out of 3	.76	.59	.64	.72	.67	4.62	5.50a
2+ out of 3	.52	1.0	1.0	.69	.77	-- b	15.49**
3 out of 3	.10	1.0	1.0	.54	.56	-- b	2.20
Trails B	.29	.91	.75	.57	.60	4.00	2.69
Digit Span (SS)	.38	.95	.89	.62	.67	12.92	7.31*
CPT Index	.67	.73	.70	.70	.70	5.33	6.70*

Note. PPP = Positive Predictive Power; NPP = Negative Predictive Power; OPP = Overall Predictive Power; ODDS = Odds-Ratio. The 7-test battery includes K-ABC Hand Movements, Stroop Color-Word Association Test, COWAT (FAS), Trails B, the Arithmetic and Digit Span Subtests of the WISC-III, and Conners CPT Overall Index. The 3-test battery includes Trails B, Digit Span, and CPT.
$^*p < .01$, $^{**}p < .001$, $^{a}\ p < .05$
[b] Odds-ratio undefined because no non-ADHD children met criteria

ered individually, only the CPT could reliably predict ADHD on its own, OPP = .70, with sensitivity, specificity, positive predictive power, and negative predictive power all between .67 and .73. Overall findings for Digit Span were also significant, OPP = .67, but sensitivity was particularly weak (.38), with the majority of ADHD participants receiving an unimpaired score. The overall predictive power of Trails B (.60) approached significance, however, sensitivity was low (.29).

Previous studies have suggested that although children with pure ADHD do display cognitive dysfunction, it may be more pronounced for those who carry the comorbid diagnosis of ODD or CD (Moffitt and Silva, 1988; Oosterlaan and Sargeant, 1996). In the present study, ADHD participants were reflective of prevalence rates described in the general population, with 76 percent ($N = 16$) meeting diagnostic criteria for either ODD or CD. A comparison of ADHD children with the comorbid diagnosis of either ODD or CD to the control group revealed patterns of impairment similar to those of the larger ADHD group. Compared to controls, ADHD children with the comorbid diagnosis of ODD/CD demonstrated impairment on the CPT overall index, $F(1, 36) = 21.82$, $p < .001$. Findings approached significance for Digit Span ($F(1, 36) = 6.68$, $p < .05$), and for Trails B ($F(1, 36) = 3.25$, $p = .08$). Classification analyses comparing the ADHD comorbid sample to the control sample are presented in Table 4. The 2 out of 7 battery approach again yielded the strongest predictive power (OPP = .82) and was slightly higher than that found in the entire sample. Predictive power of each of the three tests that demonstrated significant or near significant mean differences was also consistently though not substantially higher in this comorbid sample.

DISCUSSION

The present study sought to determine the predictive power of several tests of frontal/executive function used in combination in classifying children with ADHD. Of the seven tests, group differences emerged only for the CPT, Digit Span, and the Trailmaking Test Part B, and only the CPT provided adequate predictive power. Similar to findings recently reported by Doyle et al. (2000), a battery approach to analysis provided additional predictive power, although with

Table 4. Predictive Power of a Seven-Test Battery, Three-Test Battery, and Three Individual Tests for Children with Comorbid ADHD and ODD/CD.

	Sensitivity	Specificity	PPP	NPP	OPP	ODDS	$\chi 2$
1+ out of 7	.81	.45	.52	.77	.61	3.61	2.94
2+ out of 7	.69	.91	.85	.80	.82	22.00	14.65**
3+ out of 7	.31	1.0	1.0	.67	.71	–[b]	7.92*
4+ out of 7	.13	1.0	1.0	.61	.63	–[b]	2.90
1+ out of 3	.81	.59	.59	.81	.68	6.26	6.18[a]
2+ out of 3	.56	1.0	1.0	.76	.82	–[b]	16.22**
3 out of 3	.13	1.0	1.0	.61	.63	–[b]	2.90
Trails B	.31	.91	.71	.65	.66	4.55	3.03
Digit Span (SS)	.44	.95	.88	.70	.74	16.33	8.57*
CPT Index	.75	.73	.67	.80	.74	8.00	8.46*

Note. PPP = Positive Predictive Power; NPP = Negative Predictive Power; OPP = Overall Predictive Power; ODDS = Odds-Ratio. The 7-test battery includes K-ABC Hand Movements, Stroop Color-Word Association Test, COWAT (FAS), Trails B, the Arithmetic and Digit Span Subtests of the WISC-III, and Conners CPT Overall Index. The 3-test battery includes Trails B, Digit Span, and CPT.
*$p < .01$, **$p < .001$, [a] $p < .05$
[b] Odds-ratio undefined because no non-ADHD children met criteria

less accuracy than has been previously reported with adults (Lovejoy et al., 1999).

Impairment in performance on two or more of the seven neuropsychological tests in the present study offered the greatest overall predictive power (.77) of the various battery criteria. This battery offered strong specificity (.91), negative predictive power (.71), and positive predictive power (.87), as well as moderate sensitivity (.62). As has been found in previous studies utilizing other frontal/executive tests, an impaired score on the battery was generally indicative of ADHD, but an unimpaired score could not reliably rule out the disorder (Grodzinsky & Barkley, 1999; Lovejoy et al., 1999; Matier-Sharma, Perachio, Newcorn, Sharma, & Halperin, 1995). Although sensitivity was modest using the battery approach, it was considerably better than sensitivity found in this and other (Grodinsky & Barkley, 1998) studies of individual neuropsychological tests. Furthermore, overall predictive power using the 2 of 7 approach was comparable to the highest predictive power reported by Doyle et al. (2000), which involved using a 1 standard deviation cutoff and 2 of 7 impaired tests. However, the present study found higher negative predictive power, positive predictive power, and specificity than Doyle et al. (2000) found using this approach.

The specificity, negative predictive power, and positive predictive power found using this battery approach with children was comparable to that found with adults, however the overall predictive power and sensitivity were not as strong as reported by Lovejoy et al. (1999). It is possible that the adult participants examined by Lovejoy et al. represent a more neurologically impaired subgroup, with symptoms persisting into adulthood, than would be measured in a childhood ADHD sample. In addition, participants in the Lovejoy et al. study were selected on the basis of positive response to stimulant medication in order to obtain a sample of individuals with presumedly biologically-based ADHD.

Consistent with the majority of earlier studies, the CPT provided good predictive power, and was able to classify group membership with 70 percent accuracy. As suggested by the test developers (Conners, 1995), the overall index of the CPT provided stronger predictive power than any individual test dimension. Interestingly, the CPT overall index combines a number of indices of CPT performance and is similar in some ways to a battery approach using multiple tests. Considered separately, the differences between groups in number of hits and omission errors on the CPT approached significance, with the ADHD group displaying the weaker performance on both dimensions, consistent with previous research. Although differences in the number of commission errors approached significance between groups, subsequent analysis revealed that controls had, on average, made more commission errors than did their ADHD counterparts. It is unclear why the control participants might have made more commission errors than ADHD participants on the Conners' CPT, but this finding raises concern about the use of this dimension for both research and diagnostic purposes.

There were differences in Digit Span performance between experimental groups that approached significance, and its overall classification rate of ADHD status was modest at 67 percent. Sensitivity was particularly low for this test (.38), while specificity (.95) and positive predictive power (.89) were more impressive. Interestingly, separate analysis of group performance on the Digits Forward and Digits Backwards portions of this test revealed group differences approaching significance on the Digits Forward task only. This finding stands in contrast with that previously reported for both children (Mariani & Barkley, 1997) and adults (Jenkins et al., 1996), where impairment on the Digits Backwards task was more pronounced for those with ADHD. When an individual performs more strongly on Digits Backwards than Digits Forward, this disparity is generally believed to reflect not a lack in ability, but a lack of effort on the simpler task (Lezak, 1995). The ADHD participants in the present study may have been more strongly motivated by the additional manipulation required to perform Digits Backwards, when they performed comparably to control group participants. This finding would lend support to a behavioral disinhibition view of

ADHD, as intrinsic interest in the task may have enhanced performance (Barkley, 1998).

The absence of group differences on the COWAT and the Arithmetic subtest of the WISC-III is consistent with the mixed findings reported on these tests. More surprising was a lack of significant findings for the Hand Movements subtest of the K-ABC and the interference portion of the Stroop Color-Word Association Test. Each of these tests has received considerable attention in the literature for its reliability in classifying children with ADHD (Grodzinsky & Barkley, 1999). Although limited in number, previous ADHD studies utilizing the Hand Movements test have consistently reported significant group differences (Breen, 1989; Grodzinsky & Diamond, 1992; Mariani and Barkley, 1997). While less consistent, ADHD studies which have found significant group differences on the Stroop interference trial (Boucagnani & Jones, 1989; Carlson et al., 1986; Carter et al., 1995; Gorenstein et al., 1989; Grodzinsky & Barkley, 1999; Grodzinsky & Diamond, 1992; Hopkins et al., 1979; Seidman et al., 1997) greatly outnumber those which have not (Cohen et al., 1972; Lufi, Cohen, and Parish-Plass, 1990). It is possible that the failure to find differences on these measures in the present study was due to low power. The effect sizes for these measures were small and a larger sample may be needed to detect effects of this magnitude.

Although a battery approach to neuropsychological assessment of ADHD appears to enhance the predictive power of these tests, this study suggests that these measures fail to detect a large percentage of children with ADHD. One explanation for the modest predictive power of frontal/executive tests in classifying ADHD is the likely heterogeneity of the ADHD population. Along with chronic inconsistencies in ADHD diagnosis (National Institutes of Health, 1998), the current approach to diagnosis may be capturing children with similar symptoms, but vastly different etiologies. Neurological and genetic factors are believed to be the strongest contributors to ADHD, but environmental factors may exacerbate its course (Barkley, 1998). Some children diagnosed with ADHD may be more biologically influenced than others, and this may translate into greater neuropsychological impairment. An alternative explanation is that tests of frontal/executive functioning offer more accurate discrimination of ADHD than has been previously afforded by clinical interview; it is conceivable that neuropsychological tests offer fine prediction, and it is diagnostic interviews which are less accurate.

Another possible source of variance in both the present study and the general ADHD literature involves comorbidity of diagnoses. Since ADHD coupled with a diagnosed learning disability (LD) has been linked to poorer performance on tests of frontal/executive function than an ADHD diagnosis alone (Lazar & Frank, 1998), the present study excluded children who carried a known diagnosis of LD. Less is known about the performance of ADHD children who carry a comorbid diagnosis of either Oppositional Defiant Disorder (ODD) or Conduct Disorder (CD) when compared to those who do not. While neuropsychological impairment has been linked to CD in the literature, most studies have not reported whether the children being studied also carried an ADHD diagnosis. The limited number of studies discriminating pure ADHD from pure ODD/CD or ADHD with comorbid ODD/CD diagnoses have produced mixed results. Some have found a stronger link between executive functioning and ADHD symptoms than between executive functioning and CD symptoms (Aronowitz, Liebowitz, Hollander and Fazzini, 1994; Halperin et al., 1995; McBurnett et al., 1993; Nigg, Hinshaw, Carte, and Treuting, 1998; Oosterlaan, Logan, and Sergeant, 1998). Other studies have suggested that although children with pure ADHD do display cognitive dysfunction, it may be more pronounced for those who carry the diagnosis of ODD or CD (Moffitt and Silva, 1988; Oosterlaan and Sargeant, 1996).

In the present study, ADHD children with comorbid ODD/CD demonstrated patterns of impairment similar to those of the larger ADHD group. Classification rates for specific tests as well as for the battery approach were consistently, though not substantially, higher in identifying this comorbid sample. The overall predictive power of the 2 of 7 battery approach was

.82, with a respectable .69 sensitivity. This preliminary analysis suggests that comorbidity of ODD or CD may introduce additional variance in the study of the relationship between frontal/executive function and ADHD. It is possible that the presence of oppositional or antisocial characteristics may be associated with additional neuropsychological impairment. Future studies should more closely consider these comorbid diagnoses as well as classification rates of the various neuropsychological instruments being utilized.

Given the noted discrepancies with previous research findings, methodological limitations of the present study should be considered. While sample sizes were not greatly discrepant from those used by Lovejoy et al. (1999), the relatively low power of the present study allows for the finding of large effect sizes only. A larger sample size might reveal group differences on tests which have discriminated in past research. Generalizability to the general ADHD population is further limited by the use of male children aged six to 12, as well as the exclusion of children with diagnosed learning disorders (LD). Comparisons between ADHD children, LD children who do not have ADHD, and a non-disordered control group might have determined whether the impairments identified in the present study were specific to ADHD. Matier-Sharma et al. (1995) highlighted the additional difficulty of discriminating between ADHD children and those with other psychiatric or neuropsychological disorders within a clinic population, a more appropriate control group for assessing the predictive power of these combined measures. Similarly, the control group in the present study consisted of families of higher than average education and income; future studies are needed using more representative control groups.

Another limitation of the present study is that the order of administration of the neuropsychological test instruments was not randomized, but fixed. Although some previous studies report using a randomized test design (e.g., Barkley et al., 1992; Grodzinsky & Barkley, 1999; Grodzinsky & Diamond, 1992) others report a fixed order of testing (Seidman et al., 1997), and there were no previous reports of an effect for testing order. In the present study, there were no significant group findings for the first two tests administered, Hand Movements and the Stroop, as had been demonstrated in the majority of previous studies. More importantly, the two tests for which group differences were significant (Digit Span and CPT) were the last two tests administered to each participant. It is possible that differences in performance for ADHD children emerged at the end of the one-hour testing session, and order effects should be considered a potential influence in future studies of ADHD functioning.

Despite the limitations of the present study, it adds to the existing literature in a number of ways. While less powerful than has been demonstrated with adults (Lovejoy et al., 1999), a battery approach shows considerable promise. This study also examined not only mean group differences, but the predictive validity of several neuropsychological tests in classifying children with ADHD. Although levels of sensitivity were often low for both individual tests and the battery approach, positive predictive power was often high, indicating that impairment using the battery approach was often a good predictor of ADHD. However, low negative predictive power of the battery approach indicated that normal scores should be interpreted cautiously as they are only moderately predictive of an absence of the disorder. Continued study is needed to better understand the source of the variability in findings concerning neuropsychological testing of ADHD children and to explore further the utility of battery approaches in assessing ADHD.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, D.C.

Anastopoulos, A.D., Spisto, M.A., & Maher, M.C. (1994). The WISC-III freedom from distractibility factor: Its utility in identifying children with attention deficit hyperactivity disorder. *Psychological Assessment, 6,* 368–371.

Aronowitz, B., Liebowitz, M.R., Hollander, E., Fazzini, E., et al. (1994). Neuropsychiatric and neuropsychological findings in conduct disorder and attention-deficit hyperactivity disorder. *Journal of Neuropsychiatry and Clinical Neurosciences, 6*, 245–249.

Barkley, R.A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65–94.

Barkley, R.A. (1998). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment (2nd ed.).* New York: Guilford.

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*, 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*, 163–188.

Benton, A.L., & Hamsher, K. (1978). *Multilingual Aphasia Examination* (Manual, revised). Iowa City, IA: University of Iowa.

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.

Breen, M.J. (1981). Comparison of the Wechsler intelligence scale for children (revised) and the Peabody picture vocabulary test (revised) for a referred population. *Psychological Reports, 49*, 717–718.

Breen, M.J. (1989). Cognitive and behavioral differences in ADHD boys and girls. *Journal of Child Psychology and Psychiatry, 30*, 711–716.

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.

Carter, C.S, Krener, P., Chaderjian, M., Northcutt, C., & Wolfe, V. (1995). Abnormal processing of irrelevant information in attention deficit hyperactivity disorder. *Psychiatry Research, 56*, 59–70.

Castellanos, F.X. (1997). Neuroimaging of attention-deficit hyperactivity disorder. *Child and Adolescent Psychiatric Clinics of North America, 6*, 383–411.

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.

Cohen, M., Becker, M.G., & Campbell, R. (1990). Relationships among four measures of assessment of children with attention deficit-hyperactivity disorder. *Journal of School Psychology, 28*, 189–202.

Cohen, N.J., Weiss, G., & Minde, K. (1972). Cognitive styles in adolescents previously diagnosed as hyperactive. *Journal of Child Psychology and Psychiatry, 13*, 203–209.

Conners, C.K. (1995). *Conners' continuous performance test: User's manual.* North Tonawanda, NY: Multi-Health Systems.

Doyle, A. E., Biederman, J., Seidman, L. J., Weber, W., & Faraone, S. V. (2000). Diagnostic efficiency of neuropsychological test scores for discriminating boys with and without ADHD. *Journal of Consulting and Clinical Psychology, 68,* 477–488.

Dunn, L.M., & Dunn, L.M. (1981). *Peabody picture vocabulary test – revised.* Circle Pines, MN: American Guidance Service.

Evans, R.W. (1987). Psychostimulant pharmacology in traumatic brain injury. *Journal of Head Trauma and Rehabilitation, 2*, 29–33.

Felton, R.H., Wood, F.B., Brown, I.S., & Campbell, S.K. (1987). Separate verbal memory and naming deficits in attention deficit disorder and reading disability. *Brain and Language, 31,* 171–184.

Fischer, M., Barkley, R.A., Edelbrock, C.S., & Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria: II. Academic, attentional, and neuropsychological status. *Journal of Consulting and Clinical Psychology, 58*, 580–588.

Gaddes, W.H., & Crockett, D.J. (1975). The Spreen-Benton Aphasia Tests: Normative data as a measure of normal language development. *Brain and Language, 2*, 257–280.

Gorenstein, E.E., Mammato, C.A., & Sandy, J.M. (1989). Performance of inattentive-overactive children on selected measures of prefrontal-type function. *Journal of Clinical Psychology, 45*, 619–632.

Grodzinsky, G.M., & Barkley, R.A. (1999). The predictive power of frontal lobe tests in the diagnosis of attention deficit hyperactivity disorder. *Clinical Neuropsychologist, 13,* 12–21.

Grodzinsky, G.M., & Diamond, R. (1992). Frontal lobe functioning in boys with attention-deficit hyperactivity disorder. *Developmental Neuropsychology, 8*, 427–445.

Halperin, J.M., Newcorn, J.H., Matier, K., Bedi, G., Hall, S., & Sharma, V. (1995). Impulsivity and the initiation of fights in children with disruptive behavior disorders. *Journal of Child Psychology and Psychiatry, 36*, 1199–1211.

Hopkins, J., Perlman, T., Hechtman, L., & Weiss, G. (1979). Cognitive style in adults originally diagnosed as hyperactives. *Journal of Child Psychology and Psychiatry, 20*, 209–216.

Jenkins, M., Malloy, P., Cohen, R., Salloway, S., Neeper, R., Penn, J., & Chang, K. (1996). Attentional and learning dysfunction among adults with a history of childhood ADHD. Paper presented at the 19th annual meeting of the International Neuropsychological Society, Veldhoven, Netherlands.

Johnson, M. A. (1991). *Cognitive differences in attention deficit disorder: Differences betweeen attention deficit disorder with and without hyperactiv-*

ity. Unpublished doctoral dissertation. Fielding Institute Los Angeles, CA.

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

Lazar, J.W. & Frank, Y. (1998). Frontal systems dysfunction in children with attention-deficit/hyperactivity disorder and learning disabilities. *The Journal of Neuropsychiatry and Clinical Neurosciences, 10*, 160–167.

Lezak, M.D. (1995). *Neuropsychological assessment (3rd ed.).* New York: Oxford University Press.

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, D.W., Ball, J.D., Keats, M., Stutts, M.A., Spain, E.H., Janda, L. & Janusz, J. (1999). Neuropsychological assessment of adults with attention deficit hyperactivity disorder (ADHD): Diagnostic classification estimates for measures of frontal lobe/executive functioning. *Journal of the International Neuropsychological Society, 5,* 222–233.

Lufi, D., Cohen, A., & Parish-Plass, J. (1990). Identifying attention deficit hyperactive disorder with the WISC-R and the Stroop color and word test. *Psychology in the Schools, 27*, 28–34.

Luria, A.R. (1966). *Higher cortical functions in man.* New York: Basic Books.

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*, 118–127.

Mariani, M.A. & Barkley, R.A. (1997). Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. *Developmental Neuropsychology, 13*, 111–129.

Moffitt, T.E., & Silva, P.A. (1988). Self-reported delinquency, neuropsychological deficit, and history of attention deficit disorder. *Journal of Abnormal Child Psychology, 16*, 553–569.

McBurnett, K., Harris, S.M., Swanson, J.M., Pfiffner, L.J., Tamm, L., & Freeland, D. (1993). Neuropsychological and psychophysiological differentiation of inattention/overactivity and aggression/defiance symptom groups. *Journal of Clinical Child Psychology, 22*, 165–171.

McGee, R., Williams, S., Moffitt, T., & Anderson, J. (1989). A comparison of 13-year old boys with attention deficit and/or reading disorder on neuropsychological measures. *Journal of Abnormal Child Psychology, 17*, 37–53.

Moffitt, T. E., & Silva, P. A. (1988). Self-reported delinquency, neuropsychological deficit, and history of attention deficit disorder. *Journal of Abnormal Child Psychology, 16,* 553–569.

National Institutes of Health (1998). NIH Consensus Statement Online. Diagnosis and treatment of Attention Deficit Hyperactivity Disorder. Nov 16–18. 16 (2): 1–37.

Nigg, J.T., Hinshaw, S.P., Carte, E.T., & Treuting, J.J. (1998). Neuropsychological correlates of childhood attention-deficit/hyperactivity disorder: Explainable by comorbid disruptive behavior or reading problems? *Journal of Abnormal Psychology, 107*, 468–480.

Oosterlaan, J., Logan, G.D., & Sergeant, J.A. (1998). Response inhibition in AD/HD, CD, comorbid AD/HD + CD, anxious, and control children: A meta-analysis of studies with the stop task. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 39*, 411–425.

Oosterlaan, J. & Sergeant, J.A. (1996). Inhibition in ADHD, aggressive, and anxious children: A biologically based model of child psychopathology. *Journal of Abnormal Child Psychology, 24*, 19–36.

Ostrander, R., Wienfurt, K.P., Yarnold, P.R., & August, G.J. (1998). Diagnosing attention deficit disorders with the Behavioral Assessment Schedule for Children and the Child Behavior Checklist: Test and construct validity analyses using optimal discriminant classification trees. *Journal of Consulting and Clinical Psychology, 66*, 660–672.

Posner, M.I. & Dehaene, S. (1994). Attentional networks. *Trends in Neuroscience, 17*, 75–79.

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

Reitan, R.M., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery.* Tuscon: Neuropsychological Press.

Reynolds, C.R., & Kamphaus, R.W. (1992). *Behavior assessment system for children.* Circle Pines, MN: American Guidance Service.

Riccio, C.A., Cohen, M.J., Hall, J. & Ross, C.M. (1997). The third and fourth factors of the WISC-III: What they don't measure. *Journal of Psychoeducational Assessment, 15*, 27–39.

Sattler, J.M. (1992). *Assessment of children: Revised and updated third edition.* San Diego: Jerome M. Sattler, Publisher, Inc.

Seidman, L.J., Biederman, J., Faraone, S.V., Weber, W., & Ouellette, C. (1997). Toward defining a neuropsychology of attention deficit-hyperactivity disorder: Performance of children and adolescents from a clinically referred sample. *Journal of Consulting and Clinical Psychology, 65*, 150–160.

Semrud-Clikeman, M., Hynd, G.W., Lorys, A.R., & Lahey B.B. (1994). Differential diagnosis of children with ADHD and ADHD/ with co-occurring conduct disorder. *School Psychology International, 14,* 361–370.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., Schwab-Stone, M. E. (2000). NIMH Diagnostic

Interview Schedule for Children Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 28–38.

Shue, K.L., & Douglas, V.I. (1992). Attention deficit hyperactivity disorder and the frontal lobe syndrome. *Brain and Cognition, 20*, 102–124.

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

Wechsler, D. (1991). *Manual for the Wechsler intelligence scale for children – third edition.* San Antonio: The Psychological Corporation.