Research Article https://doi.org/10.12973/ejper.4.1.37

European Journal of Psychology and Educational Research

Volume 4, Issue 1, 37 - 49.

ISSN: 2589-949X
https://www.ejper.com

Construct Validity and Diagnostic Utility of the Woodcock Johnson Tests of Cognitive Abilities and Clinical Clusters for Children with Attention Deficit/ Hyperactivity Disorder: A Preliminary Investigation

Linda A. Reddy[*]
Rutgers University, USA

Alexander Alperin
Rutgers University, USA

Adam Lekwa
Rutgers University, USA

Received: February 4, 2021 ▪ Revised: May 19, 2021 ▪ Accepted: June 1, 2021

Abstract: The diagnostic utility of the Woodcock- Johnson III Tests of Cognitive Abilities and Clinical Clusters was assessed in a sample of 52 children (26 Attention Deficit/Hyperactivity Disordered (ADHD) and 26 matched controls). Multivariate analysis of variance followed by post-hoc testing and d-ratios yielded some statistically significant and clinically meaningful differences between groups on the Cognitive Fluency Cluster and the Tests of Auditory Attention, and Rapid Picture Naming. Discriminant function analyses indicated that the WJ III COG Tests collectively classified 80.77% of the sample correctly (76.92% of controls and 84.62% of children with ADHD correctly identified). The Auditory Attention and Rapid Picture Naming tests were found to make the most significant contribution overall to the discriminant function. Using a cut-score of 85, the WJ-III COG Clinical clusters and subtests examined in this study offered fair to weak diagnostic utility based on indices of sensitivity, specificity, positive and negative predictive power, as well as results of Receiver Operating Characteristic curve analyses. Implications for research and practice are outlined.

Keywords: *Assessment, attention deficit hyperactivity disorder, Cattell-Horn-Carroll.*

To cite this article: Reddy, L.A., Alperin, A., & Lekwa, A. (2020). Construct validity and diagnostic utility of the Woodcock Johnson tests of cognitive abilities and clinical clusters for children with attention deficit/ hyperactivity disorder: A preliminary investigation. *European Journal of Psychology and Educational Research*, *4*(1), 37-49. https://doi.org/10.12973/ejper.4.1.37

Introduction

Attention Deficit/Hyperactivity Disorder (ADHD) is one of the most commonly diagnosed child clinical disorders (e.g., Kempton et al., 1999; Tannock, 1998), representing approximately 3 to 5% of school-aged children, one to two children per classroom (e.g., DSM-IV-TR; APA, 2000; Tannock, 1998). ADHD is viewed as a complex neurocognitive disorder with impairments in alertness, arousal, selective or focused attention, sustained attention to task, and persistence of effort or vigilance (e.g., Barkley, 1997; Kempton et al., 1999; Reddy et al., 2013). The diagnosis of ADHD is typically based on a comprehensive multi-method assessment that includes cognitive, neuropsychological, achievement, and/or behavioral measures.

Many researchers assert that the behavioral manifestations of ADHD often observed by teachers and parents arise from deficits in executive functioning (Barkley, 1997; Castellanos & Tannock, 2002; Kempton et al., 1999; Reddy & Hale, 2007). Executive functioning is a collective term used to describe specific higher-level, more complex thought processes, which are considered necessary to become a successful independent learner. Factors that underlie executive functioning are working memory, organization of attentional resources, and inhibition of inappropriate responses (Kempton et al., 1999).

Closely related to executive functioning, is Gray's (1994) neuropsychological theory, which posits that behavior is governed by the behavioral activation system (BAS) and the behavioral inhibition system (BIS). The BAS connects the hypothalamus to other brain regions and motivates behavior through our desire for pleasure, satisfaction, excitement,

---

[*] **Corresponding author:**
Linda A. Reddy, Rutgers University. New Brunswick, NJ, USA. ✉ lreddy@rutgers.edu

and well-being. The BIS is responsible for inhibiting a person's actions when the behavior might be punished or not productive. For instance, the BAS might motivate a person to drive 90 mph, but the BIS would cause that person to slow down in order to avoid trouble from the police (Weis, 2017).

Barkley's (1997) theoretical model expands on Gray's (1994) theory in explaining ADHD as a deficit in executive function or the BIS, which can be illustrated in a child with ADHD as an inability to delay response to an event, interrupt an on-going response when feedback is provided, and sustain goal-directed behavior when competing stimuli are present. Barkley (1997) proposed that executive function is expressed in the areas of nonverbal working memory, verbal memory, self-regulation of emotion and motivation, and reconstitution (i.e., behavioral analysis and synthesis). The domains of executive functioning that are indirect indices of fronto-striatal systems include (a) vigilance and distractibility, (b) planning and organizing, (c) response inhibition, (d) set shifting and categorization, (e) selective attention, (f) visual scanning, and (g) verbal learning and memory (e.g., Doyle et al., 2000; Hale et al., 2010; Reddy et al., 2013).

Diagnosis of ADHD in children is often based on neuropsychological assessments. Traditional neuropsychological measures such as the Delis Kaplan Executive Functioning System (DKEFS; Delis et al., 2001), Rey-Osterrieth Complex Figure Test (ROCF; Meyers & Meyers, 1995), and Wisconsin Card Sorting Test (WCST; Heaton et al., 1993) purport to measure aspects of executive functioning among children and/or adults with ADHD. However, most school psychologists rely on traditional cognitive assessment measures and behavior rating scales when assessing children with ADHD-related symptoms (Hale et al., 2010; Reddy et al., 2013). Moreover, approximately 75% of school psychologists nationwide are Masters-level practitioners who have limited or no training in neuropsychological assessment (Koonce, 2007). In a 2017 National Survey of Assessment Practices in School Psychology, Benson et al. (Benson et al., 2019) revealed that most school psychologists do not administer neuropsychological assessments, but often administers cognitive assessments for a range of referrals including children with ADHD-related symptoms.

Cognitive theory and empirical evidence suggest relationships between data from cognitive assessments and certain aspects of ADHD. Consistent with Barkley's theoretical model of children with ADHD having greater difficulties with executive functioning is the Planning, Attention, Simultaneous, and Successive Theory (PASS; Bartolucci & Batini, 2020; Das et al., 1994). The PASS theory emphasizes that children with ADHD are more impulsive in their cognitive processing, which then has a negative effect on their planning and processing. Studies of the DAS-Naglieri Cognitive Assessment System (CAS; Naglieri & Das, 1997; Naglieri et al., 2014), which is based off the PASS theory, shows that children with ADHD typically have the lowest performance on Planning and Attention, but are more comparable to normative peers on the Simultaneous and Successive processing scores (Canivez & Gaboury, 2016; Crawford, 2002; Naglieri & Das, 1997; Naglieri et al., 2003; Naglieri et al., 2004; Paolitto, 1999; Pottinger, 2002; Van Luit et al., 2005). Similarly, Goldstein and Naglieri (2015) conducted research on the Cognitive Assessment System – Second Edition (CAS2; Naglieri et al., 2014) and found that a sample of 82 children with ADHD earned low scores on the Planning and Attention scales.

Woodcock Johnson III Tests of Cognitive Abilities and Reported Validity Concerns

A well-known and popular cognitive assessment instrument used in schools is the Woodcock Johnson III Tests of Cognitive Abilities (WJ III COG) (Woodcock et al., 2001b). Based on the Cattell-Horn-Carroll (CHC) theory, the WJ III COG includes 20 tests that fall into three cognitive performance categories, Verbal Ability, Thinking Ability, and Cognitive Efficiency (Woodcock et al., 2001b). The WJ III COG provides a General Intellectual Ability Score (GIA) Standard or Extended Scale. The Standard GIA consists of seven tests and the Extended GIA is comprised of 14 tests. Two or more individual tests are then grouped under seven broad factors: (a) Comprehension-Knowledge (Gc), (b) Long Term Retrieval (Glr), (c) Visual Spatial Thinking (Gv), (d) Auditory Processing (Ga), (e) Fluid Reasoning (Gf), (f) Processing Speed (Gs), and (g) Short Term Memory (Gsm). In addition, five clinical clusters are derived when all 20 of the cognitive tests are administered: (a) Phonemic Awareness, (b) Working Memory, (c) Broad Attention, (d) Cognitive Fluency, and (e) Executive Processes.

Each clinical cluster consists of two or more individual tests. The Phonemic Awareness cluster includes two tests: (a) Sound Blending and (b) Incomplete Words. The Sound Blending Test measures skills related to synthesizing speech sounds. The Incomplete Words test requires knowledge and skills related to analyzing and synthesizing speech sounds.

The Executive Processes cluster includes three tests: (a) Planning, (b) Pair Cancellation, and (c) Concept Formation. The Planning Test measures mental control processes involved in determining, selecting, and applying solutions to problems. The Pair Cancellation test provides information on interference control by requiring vigilant task persistence. The Concept Formation test measures flexibility of mental set shifting.

The Broad Attention cluster provides a global measure of attention and consists of four tests: (a) Numbers Reversed, (b) Auditory Working Memory, (c) Auditory Attention, and (d) Pair Cancellation. Broad Attention is a complex and multi-faceted construct that involves focused or selective attention, vigilance or sustained attention, divided attention, and attentional capacity or working memory (Schrank et al., 2002).

The Working Memory cluster consists of two tests: (a) Numbers Reversed and (b) Auditory Working Memory. Working memory is the ability to hold information in immediate awareness while performing a mental operation on the information. The Working Memory cluster assesses Barkley's concepts of nonverbal and verbal working memory. The Cognitive Fluency cluster consists of three tests: (a) Retrieval Fluency, (b) Decision Speed and (c) Rapid Picture Naming. Cognitive Fluency measures the ease and speed by which cognitive tasks are performed.

Several studies have supported the dimensional structure of the WJ-III COG using the confirmatory factor analysis modeling (e.g., Floyd et al., 2009; Keith & Reynolds, 2012; Keith et al., 2008; Locke et al., 2011; Taub et al., 2008; Taub & McGrew, 2004; Vanderwood et al., 2002). While this validation work is noteworthy, scholars have questioned the dimensional structure of the WJ-III COG pointing out that the test authors rely exclusively on confirmatory factor analysis and did not conduct exploratory factor analysis (Dombrowski, 2013). Furthermore, Dombrowski et al. (Dombrowski, 2013; Dombrowski, 2014; Dombrowski & Watkins, 2013) have conducted independent analysis of the WJ-III COG using an exploratory bifactor analysis and demonstrated possible overfactoring, as well as a factor structure that diverges from those reported in the WJ III COG technical manual.

Many scholars have noted caution regarding use of the clinical clusters in the WJ-III COG. While the clinical clusters are attractive to practitioners as they were designed to help guide diagnostic decisions, validity evidence is relatively sparse and the efficacy of these clusters has yet to be established (McGill, 2015). Specifically, McGill (2015) found that the WJ-III COG clinical clusters had limited predictive validity in accounting for Woodcock-Johnson III Tests of Achievement score variance beyond that already accounted for by the GIA (WJ-III ACH; Woodcock et al. (2001a). Dombrowski (2014) recommends caution when interpreting lower order factors such as the clinical clusters.

However, inks between the WJ-III COG Executive Processes cluster to constructs of executive functioning have been examined in multiple studies. Carper (2003) found that there was a strong correlation between the WJ-III Executive Processes cluster and the NEPSY design fluency scale. Floyd et al. (2006) found that across all correlations between the WJ III COG clusters and D-KEFS measures, Executive Processes had the highest arithmetic average. This led Floyd et al. (2006) to conclude that, when compared to the other WJ III clinical clusters, Executive Processes had the strongest relations with measures of executive functions. Since executive functioning variables are strongly related to other cognitive ability constructs (particularly fluid reasoning), the construct of executive functioning is thought to have little discriminant validity, which makes differentiating it from general cognitive abilities challenging (Salthouse, 2005).

Despite the link between the Executive Processes cluster and constructs of executive functioning, research on the clinical utility of the WJ III COG with ADHD children is limited. For example, Penny et al. (2005) examined whether a psychoeducational perspective (Cattell-Horn-Carroll theory) on cognitive abilities is useful in understanding ADHD. Using the WJ III COG, ADHD inattention symptoms were significantly related to lower scores on tests of processing speed (Adjusted R^2=0.07) and visual spatial processing (Adjusted R^2=0.12) and ADHD hyperactive and impulsive symptoms were significantly related to visual spatial processing (Adjusted R^2=0.09). In a sample of 93 children with ADHD and 85 unmatched controls (8 to 12 years), Harrier and DeOrnellas (2005) examined group performance on WJ III COG Planning subtest and Concept Formation and Analysis Synthesis subtests to represent reconstitution. Children without ADHD performed significantly better on the Concept Formation subtest than children with ADHD (partial η^2 = .13).

In another study, Ford et al. (2003) compared the performance of 58 children with ADHD to 51 unmatched controls (age 6 to 14 years) on the WJ III COG. Statistically significant group differences on the Working Memory, Broad Attention, and Executive Processes Clusters, as well as the Concept Formation, Number Reversed, Auditory Working Memory, Planning, and Pair Cancellation tests were found. Using logistic regression analyses the WJ III COG tests of Auditory Working Memory and Planning were found as significant predictors of ADHD status. Additionally, McQuade et al. (2011) assessed 4 of the 5 clinical clusters (i.e., Executive Processes, Cognitive Fluency, Broad Attention, and Working Memory) in a sample of 272 children with and without ADHD (184 children with ADHD and 88 control children). Results indicated children in the control group performed significantly better on the Broad Attention, Cognitive Fluency, and Executive Processes clusters (Cohen's d = .61, .69, .42, respectively).

Rowland's (2013) doctoral dissertation examined two groups of children aged 6 to 12 years old. The first group consisted of 23 children diagnosed with ADHD and the second (control) group had 49 children without ADHD. MANOVA results revealed that controls scored significantly higher than the ADHD group on the two factors of long-term storage and processing speed.

Two recent studies have examined the clinical utility of the WJ IV COG with ADHD students. Spenceley et al. (2020) had 150 college students (30 students were independently diagnosed with ADHD, 60 students did not have ADHD, and 60 students did not have ADHD but were asked to simulate ADHD) complete the WJ-IV COG to differentiate feigned ADHD from independently diagnosed cases of ADHD. Results demonstrated that on measures of working memory and processing speed, there were significant differences between the two groups (feigned ADHD and actual diagnosis of ADHD) with effect sizes in the large range. In addition, authors were interested in whether college students with independent diagnoses of ADHD performed differently from nondADHD peers who were not feigning ADHD symptoms. When all students responded honestly and with adequate effort, there were no group differences found.

Hart (2020) evaluated the discrepancies in CHC factor score of 31 students diagnosed with ADHD compared to 31 controls using the WJ IV three-battery configuration. Statistically significant group differences were found revealing ADHD students exhibiting relative weaknesses in auditory processing and long-term retrieval in comparison of controls.

Present Study

While these articles provide some support for the construct validity of the WJ-III COG through demonstrating distinct group differences, findings do not provide evidence determining the use of the WJ-III COG Clinical clusters in making individual diagnostic decisions. As Canivez and Gaboury (2016) emphasize in regard to the clinical utility of an instrument, "Distinct group differences are necessary but not sufficient" (p. 520). To date, there have been no diagnostic utility investigations of the WJ II COG Tests and Clinical Clusters for children with ADHD which was the purpose of this investigation.

In this study we intended to evaluate the degree to which the WJ-III COG Clinical clusters and tests could be used to differentiate between children with and without diagnoses of ADHD using analyses of group mean differences, as well as analyses of diagnostic utility. Diagnostic utility is reflected by indexes such as overall correct classification, sensitivity, specificity, positive predictive power (PPP), negative predictive power (NPP), false positive rate, and false negative rate (e.g., Canivez & Gaboury, 2016; Kessel & Zimmerman, 1993; Landau et al., 1991; Meehl & Rosen, 1955; Milich et al., 1987). The diagnostic utility of a test is essential in determining whether a test is useful for determining whether a child can or cannot be diagnosed with ADHD.

There has been prior research regarding the diagnostic utility of instruments related to ADHD. Canivez and Sprouls (2005) found the Adjustment Scales for Children and Adolescents (ASCA; McDermott et al., 1993) to correctly classify children with ADHD at a rate of 96% and have very high diagnostic efficiency (Sensitivity = .98, Specificity = .95, PPP = .94, NPP = .98). Canivez and Gaboury (2016) investigated the CAS and found that diagnostic efficiency statistics were supportive (Sensitivity = .80, Specificity = .75, PPP = .76, NPP = .79) of the CAS's ability to be useful in diagnosing ADHD in young children. Using the Differential Ability Scale subtests (DAS; Elliott, 1990), Reddy et al. (2008) reported overall correct classification of 69% for ADHD children, and diagnostic efficiency statistics supported use of the Recall of Digits subtest (cutoffs of 1.0 below the standardization mean; sensitivity = .14, specificity = .93, PPP = .67, NPP = .54).

Building on this line of research, the present investigation serves as the first (pilot) investigation to examine the construct validity and diagnostic utility of the WJ III COG Tests and Clinical Clusters for children independently diagnosed with ADHD to a matched control sample. Specifically, this study assesses the WJ III COG and Clinical Clusters' ability to practically differentiate group and individual student differences among children with ADHD from those without this diagnosis in an outpatient clinic.

The primary research questions were: (1) Do children with ADHD perform worse on the WJ III COG Tests and Clinical Clusters than controls? and (2) What is the diagnostic utility of the WJ III COG Tests and Clinical Clusters for children with ADHD and matched controls?

Methodology

Sample

Table 1: Sample Demographic Characteristics

	Control		**ADHD**		**Total**	
Variable	**n**	**%**	**n**	**%**	**n**	**%**
Age in Months						
60-79	5	19.2	6	23.1	11	21.2
80-99	9	34.6	7	26.9	16	30.8
100-119	10	38.5	10	38.5	20	38.4
120-129	2	7.7	3	11.5	5	9.6
Gender						
Male	21	80.8	22	84.6	43	82.7
Female	5	19.2	4	15.4	9	17.3
Ethnicity						
White	21	80.8	24	92.3	45	86.5
African American	2	7.7	2	7.7	4	7.7
Hispanic/Latino	3	11.5	0	0	3	5.8

Table 1: Continued

	Control		ADHD		Total	
Variable	**n**	**%**	**n**	**%**	**n**	**%**
Mother's Level of Education						
High School	0	0	3	11.5	3	5.8
1 to 3 years of college	0	0	5	19.2	5	9.6
Bachelor's degree	4	15.4	12	46.2	16	30.8
Master's degree	10	38.5	0	0	10	19.2
Master's degree +	12	46.1	6	23.1	18	34.6
Father's Level of Education						
High School	0	0	7	26.9	7	13.5
1 to 3 years of college	0	0	4	15.4	4	7.7
Bachelor's degree	5	19.2	13	50	18	34.6
Master's degree	13	50	0	0	13	25
Master's degree +	8	30.8	2	7.7	10	19.2

Demographic information regarding the sample is presented on Table 1. The ADHD and control samples were matched on five demographic variables (age, gender, race, mother's level of education, father's level of education) similar to those used for the WJ III COG standardization sample. Chi-squared analyses indicated that the ADHD and Control subjects were comparable on gender ($\chi^2 = .13$, $df = 1$, $p = .71$), race ($\chi^2 = .35$, $df = 1$, $p = .55$), mother's level of education ($\chi^2 = 5.80$, $df = 4$, $p = .21$), and father's level of education ($\chi^2 = 3.90$, $df = 4$, $p = .42$). No significant group difference on age was found ($t = .36$, $df = 50$, $p = .71$).

The ADHD sample was recruited through a University-based Child and Adolescent ADHD research clinic. Several inclusion and exclusion criteria were used to select the ADHD sample. Children were included that: (1) received a primary diagnosis of ADHD by a pediatric neurologist, psychiatrist, and/or psychologist, (2) met the DSM-IV-TR (APA, 2000) criteria for ADHD prior to the study, (3) were 12 years and younger, and (4) were enrolled full-time in school. Children with ADHD were excluded that: (1) had parents currently separated or in the process of divorce, (2) physically or sexually abused within the past 18 months, (3) experienced other significant losses (e.g., death of a parent, sibling) in the past 12 months, and (4) diagnosed with a brain injury or seizure disorder. These exclusion criterion were used because these events have been found to contribute to symptoms (e.g., concentration, inattention) that resemble some of those found in children diagnosed with ADHD (e.g., Björkenstam et al., 2018; Brown et al., 2017).

The Structured Diagnostic Interview for Parents (SDIP; Barkley & Murphy, 2006) a DSM-IV semi-structured interview was completed with parents referred to the clinic. Approximately 58% of the ADHD sample met critera for additional psychiatric diagnoses. The sample met the DSM-IV-TR (APA, 2000) criteria for the following psychiatric diagnoses: 35% Oppositional Defiant Disorder, 31% Separation Anxiety Disorder or Anxiety Disorder Not Otherwise Specified, 15% Learning Disabled, and 8% Dysthymia. Approximately 61% of the ADHD sample received special education services under the Individual Disabilities Education Act (IDEA) or Section 504 (EveryCRSReport, 2004). The control sample did not have any psychiatric diagnoses and were not receiving any special educational services. Control subjects were selected from the standardization sample of the WJ III COG based on the five matching variables. The WJ III COG was not used for diagnostic decision making.

Instrument

WJ III COG. As noted, the WJ III COG includes 20 tests that fall into Verbal Ability, Thinking Ability, and Cognitive Efficiency that generate a General Intellectual Ability Score (GIA) for Standard and Extended batteries (Woodcock et al., 2001b). The Standard GIA consists of seven tests, the Extended GIA is comprised of 14 tests and individual tests are grouped under seven broad factors: (a) Comprehension-Knowledge (Gc), (b) Long Term Retrieval (Glr), (c) Visual Spatial Thinking (Gv), (d) Auditory Processing (Ga), (e) Fluid Reasoning (Gf), (f) Processing Speed (Gs), and (g) Short Term Memory (Gsm). Five clinical clusters are derived when all 20 of the cognitive tests are administered: (a) Phonemic Awareness, (b) Working Memory, (c) Broad Attention, (d) Cognitive Fluency, and (e) Executive Processes with each clinical cluster consists of two or more individual tests. For the purposes of this study, only the clinical clusters of Broad Attention, Cognitive Fluency, Executive Processes, and Working Memory were used; the Phonemic Awareness cluster and its two constituent tests (Incomplete Words and Sound Blending), which are not known to be associated with elements of ADHD, were excluded from analysis (please see earlier description of subscales).

The internal reliability of the scores produced by the WJ III COG is high. Split-half and Rasch reliability statistics for most of the untimed subtest scores are above .80 (range = .74 to .97), while the values for cluster scores fall mostly above .90 (range = .81 to .98). Test-retest reliability for the timed tests ranged from .70 to .96. Intercorrelations between subtests are small, ranging from .20 to .60, suggesting that they do measure different factors. Concurrent validity of the WJ III COG was .71 when compared to the WISC III and .67 when compared to the WAIS-III.

Procedure

The WJ III COG was administered by school psychology doctoral students who were trained and supervised by school psychology faculty. All graduate-level testers attended a three-hour training session focused on the theory of the WJ III COG, the utility of the test, test administration, and test response scoring. Testers successfully completed a graduate level course on cognitive assessment and had administered and scored at least two WJ III COG tests prior to the study. Test protocols were reviewed by a school psychology faculty member. Raw scores obtained were converted to standard scores by using the WJ III COG computerized scoring system (Compuscore and Profiles Program). Each raw score protocol was independently cross-checked, yielding 100% agreement. Testing was completed over two or three testing sessions.

Intercorrelations between the WJ III COG four Clinical Clusters and nine tests were computed for the entire sample to rule out multicollinearity among tests. As shown on Tables 2 and 3, the intercorrelations between the clinical clusters and tests were not consistently excessive (.90 and higher), suggesting that multicollinearity was not an issue (Tabachnick & Fidell, 1996). Furthermore, intercorrelations for the WJ III COG for this sample were comparable to those found in the published technical manual.

Table 2: Intercorrelations among the WJ III COG Clinical Clusters

Cluster	**Broad Attention**	**Cognitive Fluency**	**Executive Processing**	**Working Memory**
Broad Attention	1	0.72	0.67	0.92
Cognitive Fluency		1	0.52	0.63
Executive Processing			1	0.51
Working Memory				1

Table 3: Intercorrelations among the WJ III COG Tests

Test	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1. Auditory Attention	1	0.37	0.41	0.44	0.28	0.27	-0.14	0.52	0.36
2. Aud.Working Memory		1	0.48	0.34	0.22	0.32	0.02	0.41	0.27
3. Concept Formation			1	0.4	0.21	0.36	0.14	0.37	0.32
4. Decision Speed				1	0.47	0.74	0.02	0.53	0.26
5. Numbers Reversed					1	0.45	0.08	0.51	0.46
6. Pair Cancellation						1	0.21	0.4	0.21
7. Planning							1	-0.1	-0.13
8. Retrieval Fluency								1	0.7
9. Rapid Picture Naming									1

Parents referred to the University-based Child and Adolescent ADHD Clinic were offered free testing to assist them in learning more about their children's cognitive functioning and provided a summary of scores, as well as verbal feedback about the test results. Informed consent was obtained from all parents and assent was obtained from all children.

Data analysis

Several data analytic methods were used to assess group performance on the WJ III COG. First, descriptive statistics, correlations and Cohen *d* (effect sizes) were computed. D-ratios of .20 to .49, .50 to .69, and .70 and above signified small, medium, and large effect sizes between the samples (Cohen, 1988).

Second, multivariate analysis of variance (MANOVA) and univariate analyses were computed. Third, two direct discriminant function analyses, using a jackknife procedure for cross-validation, were carried out to assess the classification accuracy of the four Clinical Clusters and 9 Tests for ADHD status. Assumptions for both sets of analyses included freedom from outliers, as well as linear relationships, multivariate normality, and homogenous covariance estimates between the four Clinical Clusters and the nine test scores. Analysis of Mahalanobis distances revealed no significant multivariate outliers. The four Clinical Clusters, as well as the nine test scores, exhibited linear relationships (based on visual inspection of scatterplots between each variable). Moderate multicollinearity was detected between the Broad Attention and Working Memory Clinical clusters ($r = .92$). Box's M test for homogeneity of covariance estimates was insignificant at the .001 level for the Clinical scores ($\chi^2(10) = 9.94$, $p = 0.45$) and the nine subtests ($\chi^2 (45) = 62.08$, $p = .05$); see Tabachnik and Fidell, (2001). Although results of the Shapiro-Wilk test of multivariate normality suggested some deviation from normality in the four Clinical Clusters and the nine test scores, inspection of individual distributions and descriptive statistics did not reveal substantial skew in any Clinical Cluster or test score; MANOVA and discriminant function analyses are considered robust to non-normality in the absence of outliers.

Fourth, diagnostic efficiency statistics (sensitivity, specificity, negative predictive power [NPP], and positive predictive power [PPP]) and Receiver Operating Characteristic (ROC) curve analyses were computed to assess how well children with or without diagnoses of ADHD could be differentiated based on their scores on the 9 WJ-III Cog tests, or the 4 Clinical Clusters. Diagnostic efficiency statistics for each cluster or test were based on a cut score of 85 as recommended in the WJ-III Cognitive Technical Manual (Woodcock et al., 2001b). Results of the ROC analyses for each cluster or test, called Area Under the Curve (AUC) represent the ability of a particular test to aid in differentiation of children with and without ADHD across all possible cut scores.

Results

WJ-III Clinical Clusters

Group differences (research question 1)

Table 4 : Descriptive Statistics, MANOVAs, and Cohen's d-ratios for the WJ III COG Clinical Clusters and Tests

Clusters/ Tests	Control		ADHD			
	M	**SD**	**M**	**SD**	**F**	**d-ratio**
Clusters:						
Broad Attention [a]	100.15	12.24	100.57	16.24	0.01	-0.02
Cognitive Fluency	97.07	16.25	89.03	17.84	2.88 **	0.47
Executive Processing	103.8	13.53	101.8	16.55	0.22	0.13
Working Memory	101.38	12.25	104.96	16.93	0.76	-0.24
Tests:						
Auditory Attention [b]	101.53	7.67	89.53	16.13	11.73 ***	0.95
Auditory Working Memory	99	12.89	105.26	18.83	1.96	-0.38
Concept Formation	107.84	15.57	100.42	17.52	2.6	0.44
Decision Speed	101.8	16.13	101.46	17.47	0	0.02
Numbers Reversed	102.15	15.01	100.84	16.64	0.08	0.08
Pair Cancellation	96.19	11.31	99.69	13.08	1.06	-0.28
Planning	105.53	10.81	109.11	14.3	1.03	-0.28
Retrieval Fluency	104.34	13.75	97.76	17.99	2.21	0.41
Rapid Picture Naming	95.38	15.47	85.38	16.78	4.99 **	0.62

Note. [a] - λ $F(4, 47) = 2.62$; $p = .04$; [b] - λ $F(9, 42) = 4.35$; $p = .00$; * p <.05; ** p <.01; *** p <.001

As shown on Table 4, descriptive statistics suggested that the ADHD sample's mean scale scores for the four Clinical Clusters with the exception of Cognitive Fluency ($M = 89$) approximated the standardization sample ($M = 100$; $SD = 15$). Similarly, the control sample yielded scale scores approximately equal to the standardization sample.

Multivariate analysis of variances (MANOVA) tests and Cohen's *d* (effect sizes) were computed (Table 4) to determine the statistical and clinical significance between sample scale scores. The MANOVA Wilks Lambda test indicated significant group differences among the four Clinical Clusters [λ F (4, 47) = 2.62, $p = .04$]. Exploratory univariate analyses revealed that the ADHD sample scored significantly lower on Cognitive Fluency than the control sample. The Cognitive Fluency and Working Memory clusters also produced small *d*-ratios of .47 and -.24, respectively.

Diagnostic utility of the WJ-III Clinical Clusters (research question 2)

A direct discriminant function analysis was computed to assess how well the four clinical clusters and nine tests predict ADHD diagnosis. For purposes of cross-validation and bias estimation, we re-fit the discriminant function model for the Clinical clusters using a jackknife procedure; minimal shrinkage in estimates was observed between results obtained from either method (a reduction in overall classification accuracy from 69% to 65%). For exploratory purposes, the first discriminant function analysis included the four Clinical Clusters and yielded a significant $\chi^2 = 9.67$, $df = 4$, $p < .05$ (Wilks' $\lambda = .82$). The canonical correlation of .43 suggested that the discriminant function accounted for 18.49% of the variance between the samples. Structural coefficients (i.e., r) provide a simple bivariate correlation between the cluster or test and the canonical discriminant function. Huberty (1984) stated that a large r suggests that a particular cluster or test produces separation comparable to that obtained by the discriminant function. The Cognitive Fluency, and Working Memory clusters had moderate coefficients of .51, -.26, respectively; the Executive Processing and Broad Attention clusters had small coefficients of .14, and -.03, respectively. As shown in Table 5, the overall correct classification of the clinical clusters was 65.40%, yielding an increase of 15.40% in the overall hit rate when classifying ADHD children beyond chance alone. ADHD classifications based on this model had a sensitivity of .69, specificity of .62, and positive and negative predictive powers of .64 and .67, respectively.

Table 5: Latent Discrimination Analysis, ROC Analysis, and Diagnostic Efficiency Statistics

Test/Cluster	Fisher's Std. LDF Coefficients		ROC	Diagnostic Statistics [a]				
	Control	ADHD	AUC	Sensitivity	Specificity	PPV	NPV	Accuracy
Clinical Clusters			*0.65*	*0.69*	*0.62*	*0.64*	*0.67*	*0.65*
Broad Attention	-0.24	-0.28	0.55	0.19	0.96	0.83	0.54	0.58
Cognitive Fluency	0.06	0	0.63	0.46	0.69	0.6	0.56	0.58
Executive Processes	0.34	0.34	0.49	0.12	0.96	0.75	0.52	0.54
Working Memory	0.45	0.55	0.6	0.12	0.88	0.5	0.5	0.5
Tests			*0.8*	*0.85*	*0.77*	*0.79*	*0.83*	*0.81*
Auditory Attention	0.54	0.4	0.79	0.38	1	1	0.62	0.69
Auditory Working Memory	0.03	0.14	0.64	0.12	0.85	0.43	0.49	0.48
Concept Formation	0.02	-0.04	0.6	0.12	0.96	0.75	0.52	0.54
Decision Speed	-0.07	-0.07	0.52	0.12	0.85	0.43	0.49	0.48
Numbers Reversed	0.07	0.08	0.51	0.12	0.88	0.5	0.5	0.5
Pair Cancellation	0.22	0.3	0.6	0.19	0.81	0.5	0.5	0.5
Planning	0.68	0.68	0.57	0	0.92	0	0.48	0.46
Retrieval Fluency	-0.01	-0.01	0.58	0.19	0.96	0.83	0.54	0.58
Rapid Picture Naming	0.23	0.18	0.7	0.58	0.73	0.68	0.63	0.65

Note. [a] - Cut-scores of 85 were used.

The four Clinical clusters, individually, exhibited lower levels of diagnostic efficiency (Table 5). Although results varied, each cluster ranged between .49 and .63 in AUC, and between .50 and .58 in overall classification accuracy at a cut score of 85. The clusters of Broad Attention, Cognitive Fluency, and Working Memory each had low sensitivity, meaning these clusters did not tend to classify individuals in the ADHD group as having ADHD. In contrast, the sensitivity indices suggested that these three clusters performed much better at identifying those students without the disorder (identifying between .88 and .96 of the 26 matched control students *without* ADHD in the sample). Of the four clusters, Cognitive Fluency exhibited the greatest sensitivity (correctly classifying .46 of those students who had received diagnoses of ADHD), but comparatively lower specificity (.69).

The Broad Attention and Executive Processes clusters, when they did predict the presence of ADHD, tended to be slightly more accurate than the Cognitive Fluency and Working Memory clusters (with PPV indices for Broad Attention and Executive Processes at .83 and .75, versus indices at .60 and .50 for Cognitive Fluency and Working Memory). The NPV estimates (ratio of predicted non-ADHD to actual non-ADHD status) ranged between .50 and .56 for each cluster.

WJ-III Cognitive Tests

Group differences (research question 1).

For the WJ III COG Tests, the ADHD sample produced slightly lower scale scores than the control sample on Auditory Attention, Concept Formation, Incomplete Words, Numbers Reversed, Retrieval Fluency, Rapid Picture Naming, and Sound Blending (see Table 4). In general, the ADHD sample's mean scale scores for the tests approximated the standardization sample with the exception of Auditory Attention (M = 89.53) and Rapid Picture Naming (M = 85.38). Additionally, the ADHD sample yielded slightly higher scale scores on Auditory Working Memory, Pair Cancellation, and Planning than the control sample.

A MANOVA Wilks Lambda test indicated group differences for the 9 Tests (λ F [9, 42] = 4.35, $p < .01$). Univariate analyses also revealed that the ADHD sample scored significantly lower than the control sample on Auditory Attention and Rapid Picture Naming with large and medium positive d-ratios found for each (i.e., .95, and .62). Small positive d-ratios were found for Concept Formation, and Retrieval Fluency, while small negative d-ratios were found for Auditory Working Memory, Pair Cancellation, and Planning.

Diagnostic utility of WJ-III Cognitive Tests (research question 2).

The second discriminant function analysis included the 9 Tests (see Table 5). As with the Clinical clusters, we re-fit the discriminant function model for the 9 tests using a jackknife procedure; minimal shrinkage in estimates was observed between results obtained from either method (a reduction in overall classification accuracy from 83% to 81%). The Tests produced a significant χ^2 = 29.97, df = 9, $p < .001$ (Wilks' λ = .52). The canonical correlation of .70 suggested that the discriminant function accounted for approximately 49% of the variance between the samples. Moderate positive structure coefficients were observed for Auditory Attention (r = .50) and Rapid Picture Naming (r = .33). Small positive structure coefficients were observed for Concept Formation (r = .24), and Retrieval Fluency (r = .22). Small negative structure coefficients were observed for Auditory Working Memory (r = -.21), Pair Cancellation (r = -.15), and Planning

(r = -.15). Coefficients for Numbers Reversed and Decision Speed were negligible. As shown on Table 4, the overall correct classification for the tests together as a group was 80.77%, resulting in an increase of 30.77% beyond chance in the overall hit rate when classifying ADHD children. ADHD classifications based on this model had sensitivity of .85, specificity of .77, and positive and negative predictive powers of .79 and .83, respectively.

Considered individually, WJ-III COG tests demonstrated variability in diagnostic efficiency statistics at a cut-score of 85. Two of the nine individual tests, Auditory Attention and Rapid Picture Naming, exhibited diagnostic efficiency indices stronger than those of the other seven, with AUCs at or above .70, and overall classification accuracies at or above .65. Rapid Picture naming had the greatest sensitivity of all nine tests (.58), yet limited PPV and NPV (both below .70). Auditory Attention had the greatest specificity (at 1.00), limited sensitivity (.38), and nearly perfect PPV. The remaining seven individual tests had AUC values between .51 and .64, with overall accuracy estimates at approximately .51; sensitivity indices among these seven tests average .12, and did not exceed .19, although specificities averaged around .89. Indices of PPV and NPV averaged approximately .50.

Discussion

Research has found that ADHD children perform worse than controls on a wide range of cognitive and neurocognitive tests that assess aspects of executive functioning (e.g., Hart, 2020; Reddy & Hale, 2007; Reddy et al., 2013), however the WJ COG, one of the most wide used cognitive assessment batteries has very limited independent research on it use for this children and adults (Spenceley et al., 2020). In general findings from this pilot investigation highlight limited clinical utility of specific Clinical Clusters and individual tests for informing decision making for this population. For example, the clinical cluster of Cognitive Fluency and test of Rapid Picture Naming significantly discriminated the ADHD group from a matched control group.

Other Clinical Clusters expected to be related to ADHD status (i.e., Working Memory, Broad Attention, and Executive Processing) did not discriminate groups. This is in contrast to the Ford et al. (2003) study and McQuade et al. (2011) study. Ford et al. (2003) found significant differences between children with ADHD and unmatched controls on the WJ III COG for the Working Memory, Broad Attention, and Executive Processes Clusters, while McQuade et al. (2011) reported that children in the control group performed significantly better on the Broad Attention, Cognitive Fluency, and Executive Processes clusters. While both Ford et al. (2003) and McQuade et al. (2011) found significant differences for the Broad Attention and Executive Processes clusters, the participants in the present study performed similarly on both of these clusters. This is in line with Harrier and DeOrnellas (2005), who also did not find significant differences in tests of executive functioning. As mentioned earlier, Executive Processes cluster purports to measure the construct of executive functioning and Broad Attention provides a global measure of attention. The disparity in findings in this study compared to previous investigations (e.g., Ford et al., 2003; Harrier & DeOrnellas, 2005) may be due in part to sample size, sample characteristics, matched samples, diagnostic inclusion and exclusion criteria, and/or the extent of comorbidities.

The current study also examined the utility of the individual Tests of the WJ III COG. Specifically, ADHD subjects performed significantly lower than controls on the Auditory Attention and Rapid Picture Naming subtests. Rapid Picture Naming is the only test within the Cognitive Fluency Cluster that significantly discriminated between the groups. Present findings with regard to the Auditory Attention subtest are consistent with Lerner and Yasutake (2000). Additionally, Hart (2020) found students with ADHD in general exhibiting relative weaknesses in auditory processing and long-term retrieval compared to controls. While it is not surprising that a test of selective auditory attention would differentiate children with ADHD from controls, it is worth noting that other tests of attention (e.g., Numbers Reversed, Auditory Working Memory, Pair Cancellation) did not result in group differences in our study as were in Ford et al.'s (2003) study. In fact, these were found to be relative strengths in the current ADHD sample. However, the trend of lower performance found on Concept Formation in the current ADHD sample is consistent with findings from Harrier and DeOrnellas (2005) results with Inattentive and Combined subtype subjects. This trend may suggest that reconstitution is a crucial deficit in children with ADHD.

This study examined the diagnostic utility of WJ III COG scores, thereby going beyond the typical distinct group differences or discriminant validity studies. Results suggested that administering that the WJ III COG may have some promise as a cognitive measure for assisting in diagnosis of ADHD in children. While the four Clinical Clusters, together and individually, and the 9 Tests, individually, did exhibit low levels of diagnostic efficiency, the 9 tests together showed promise in helping diagnose children with ADHD. Specifically, the overall classifications for the tests taken together was 80.77%, resulting in an increase of 30.77% beyond chance in the overall hit rate when classifying ADHD children. ADHD classifications based on the 9 tests together had sensitivity of .85, specificity of .77, and positive and negative and predictive powers of .79 and .83. While shrinkage was observed when jackknifed classification was used in cross-validation, diagnostic efficiency statistics were still statistically significant and generally good. ROC analysis showed a moderate degree of diagnostic accuracy. This is comparable to the diagnostic efficiency statistics (Sensitivity = .80, Specificity = .75, PPP = .76, NPP = .79) reported by Canivez and Gaboury (2016) regarding the Das-Naglieri Cognitive Assessment System (CAS). However, the present study's results are not as high as the diagnostic efficiency statistics

(Sensitivity = .98, Specificity = .95, PPP = .94, NPP = .98) reported by Canivez and Sprouls (2005) for the Adjustment Scales for Children and Adolescents (ASCA).

In general, the WJ III COG's individual clusters and individual tests under predicted ADHD status at a cut score of .85, but were better at ruling out ADHD, with specificity estimates averaging .88, though the accuracy of the negative predications was varied and limited. In other words, the WJ III COG's individual clusters and individual tests ruled out the majority of students who did not have a diagnosis of ADHD, but also ruled out many students that actually did have a diagnosis of ADHD. Taken altogether, it appears that no single WJ-III Score appears to have utility in identifying the presence of ADHD.

This is in line with previous research from Dombrowski et al. (Dombrowski, 2013 b; Dombrowski & Watkins, 2013; Dombrowski, 2014) who questioned the structure of the WJ-III COG and demonstrated possible overfactoring. In addition, McGill (2015) questioned the WJ III COG's incremental validity. This research along with the present study signifies that the WJ-III COG might be useful in diagnosing children with ADHD when taking into account the 9 Tests described above, but due to the WJ III COG's questionable structure and construct validity, caution is advised when interpreting individual test and/or cluster scores.

Study Strengths and Limitations

The present investigation has several strengths and limitations. One strength of the present study is that several inclusion and exclusion criteria were used for the ADHD and control samples. Second, children were independently diagnosed by practitioners. Third, matched controls that did not meet criteria for specific learning disabilities and/or psychiatric disorders were included. Also, groups were carefully matched on five variables (i.e., age, gender, race, and parental education – maternal and paternal) which significantly contribute to performance on measures of cognitive ability (Lezak, 1995; Reddy et al., 2008) and are commonly used to standardize child cognitive assessment measures (Reddy & Hale, 2007; Reddy et al., 2011). The comorbidity of the ADHD sample in this study is both a strength and limitation. Fifty-eight percent of the ADHD sample has another psychiatric disorder and 62% of the sample received special education services and/or accommodations (under IDEA or section 504) for learning. These findings are consistent with what has been observed in the general ADHD population (Reale et al., 2017; Reddy & Hale, 2007; Semrud-Clikeman et al., 1994). Research has also shown that up to 44% of ADHD children have at least one other psychiatric disorder, 32% have two others, and 11% have at least three other disorders (Barkley, 1997). This breakdown is consistent with the present investigation. Thus, the comorbidity and related services characteristics in this sample is a strength for the current study and its generalizability to the ADHD population. However, comorbidity is also a limitation in that results of the study may not be attributable to children diagnosed with ADHD only.

The present investigation did not examine differences in performance among ADHD subtypes. Research has shown that ADHD subtypes represent unique and distinct disorders and can manifest different neurocognitive features (e.g., Hale, Reddy et al., 2009; Reddy & Hale, 2007; Reddy et al., 2013).

The current investigation is also limited by sample characteristics. Similar to Ford et al. (2003) and Harrier and DeOrnellas (2005), the present study lacked ethnic and geographic diversity. Therefore, the generalizability of findings to ADHD children from other ethnic, racial and geographic backgrounds is limited. Considering that many scholars and clinicians have suspected that ADHD is frequently misdiagnosed and children with auditory processing disorders are often described as easily distracted and inattentive, it is possible that some of our subjects were misdiagnosed as having ADHD (Ford-Jones, 2015; Gyldrenkaerne et al., 2014). Thus, there might be a potential confound in interpreting the statistically significant difference between the control and ADHD group for the Auditory Attention test. Finally, a limitation of the current study is the small sample size. Previous studies have found that inattentive symptoms of ADHD have been associated with poorer performances on tests of fluid reasoning or reconstitution (Harrier & DeOrnellas, 2005), planning, and working memory (Ford et al., 2003). However, the current study fails to support these findings. The performance of ADHD children was lower than controls in some of these domains, but these differences were not significant. It is possible that a larger sample (increased power) would have resulted in more significant differences between groups.

Recommendations

Although the relationship between ADHD status and variables such as auditory processing, auditory working memory, and auditory attention has received some attention (e.g., Ford et al., 2003; Hart, 2020; Lerner & Yasutake, 2000), research is needed that extends this work and critically examines the role that auditory processes plays in the neurocognitive deficits associated with ADHD. Findings from this investigation are consistent with those of Penny et al. (2005), as well as Harrier and DeOrnellas (2005), suggesting that ADHD and non-ADHD groups may differ on scores of Cognitive Fluency, but not other clinical clusters. This may be an area of critical importance in the understanding of this disorder, and should be addressed in future research and practice.

Future research on the WJ III COG should also include more diverse ADHD samples. As noted, the majority of the literature focuses on Caucasian males. Much information is not known about females with ADHD in schools. Likewise, despite the large prevalence rates of ADHD in minority populations, little is known about the neuropsychological functioning of ADHD children from ethnically diverse populations. Finally, future research using the WJ III COG should aim to distinguish subtypes of ADHD. This line of work will be valuable for informing practitioners of the possible cognitive profiles of children with specific ADHD subtypes.

Conclusion

The WJ III COG is a widely used and well researched cognitive assessment tool in schools. However, limited research exists on its clinical utility with school-aged children diagnosed with ADHD. This pilot study represents the first investigation of the use of the WJ III COG for independently diagnosed children (5 to 12 years) with ADHD and a matched control sample. Using discriminant function analyses, the WJ III COG's 11 Tests of Cognitive Abilities offer practitioners classification accuracy of 82.7%, an increase of 32.7% in the overall hit rate when classifying ADHD children, respectfully. Additionally, the WJ III COG Tests did offer practical differences (*d*-ratios) between samples (e.g., Auditory Attention, Rapid Picture Naming, Sound Blending) which provide practitioners with some clinical insights when using the WJ III COG with children at risk for or diagnosed with ADHD. As noted, additional research is warranted to maximize the clinical usefulness of WJ III COG scores for youth with ADHD.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. Text rev.). American Psychiatric Association.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94. https://doi.org/10.1037/0033-2909.121.1.65

Barkley, R. A., & Murphy, K. R. (2006). Attention–deficit hyperactivity disorder: A clinical workbook (3rd ed.). The Guilford Press.

Bartolucci, M., & Batini, F. (2020). Reading aloud narrative material as a means for the student's cognitive empowerment. *Mind, Brain, and Education*, *14*(3), 235-242. https://doi.org/10.1111/mbe.12241

Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefer, S. A., & Morgan, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 national survey. *Journal of School Psychology*, *72*, 29-48. https://doi.org/10.1016/j.jsp.2018.12.004

Björkenstam, E., Björkenstam, C., Jablonska, B., & Kosidou, K. (2018). Cumulative exposure to childhood adversity, and treated attention deficit/hyperactivity disorder: a cohort study of 543 650 adolescents and young adults in Sweden. *Psychological medicine, 48*(3), 498-507. https://doi.org/10.1017/S0033291717001933

Brown, N. M., Brown, S. N., Briggs, R. D., German, M., Belamarich, P. F., & Oyeku, S. O. (2017). Associations between adverse childhood experiences and ADHD diagnosis and severity. *Academic pediatrics, 17*(4), 349-355. https://doi.org/10.1016/j.acap.2016.08.013

Canivez, G. L., & Gaboury, A. R. (2016). Construct validity and diagnostic utility of the cognitive assessment system for ADHD. *Journal of attention disorders, 20*(6), 519-526. https://doi.org/10.1177/1087054713489021

Canivez, G. L., & Sprouls, K. (2005). Assessing the construct validity of the Adjustment Scales for Children and Adolescents. *Journal of Psychoeducational Assessment, 23*(1), 3-14. https://doi.org/10.1177/073428290502300101

Carper, M. S. (2003). *A concurrent validity study comparing executive functioning of the Woodcock-Johnson III Tests of Cognitive Ability and the NEPSY* [Unpublished doctoral dissertation]. Texas Woman's University.

Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience*, *3*(8), 617-628.

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Academic Press.

Crawford, E. N. (2002). Profiles for exceptional samples on the Cognitive Assessment System using configural frequency analysis [Unpublished doctoral dissertation]. University of Northern Colorado.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence.* Allyn & Bacon.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis Kaplan executive function system*. The Psychological Corporation.

Dombrowski, S. C. (2013). Investigating the structure of the WJ-III Cognitive at school age. *School Psychology Quarterly, 28*(8), 154-169. https://doi.org/10.1037/spq0000010

Dombrowski, S. C. (2014). Exploratory bifactor analysis of the WJ-III Cognitive in adulthood via the Schmid-Leiman procedure. *Journal of Psychoeducational Assessment, 32*(4), 330 -341. https://doi.org/10.1177/0734282913508243

Dombrowski, S. C., & Watkins, M. W. (2013). Exploratory and higher order factor analysis of the WJ-III full test battery: A school-aged analysis. *Psychological Assessment, 25*(2)*,* 442-455. https://doi.org/10.1037/a00311335

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 attention deficit/hyperactivity disorder. *Journal of Consulting and Clinical Psychology, 68*(3), 477-488. https://doi.org/10.1037/0022-006X.68.3.477

Elliott, C. D. (1990). *Differential Ability Scales.* The Psychological Corporation.

Floyd, R. G., McCormack, A. C., Ingram, E. L., Davis, A. E., Bergeron, R., & Hamilton, G. (2006). Relations between the Woodcock-Johnson III clinical clusters and measures of executive functions from the Delis-Kaplan Executive Function System. *Journal of Psychoeducational Assessment*, *24*(4), 303-317. https://doi.org/10.1177/0734282906287823

Floyd, R. G., McGrew, K. S., Barry, A., Rafael, F., & Rogers, J. (2009). General and specific effects on Cattell–Horn–Carroll broad ability composites: Analysis of the Woodcock–Johnson III normative update Cattell–Horn–Carroll factor clusters across development. *School Psychology Review, 38*(2)*,* 249 –265. https://doi.org/10.1080/02796015.2009.12087836

Ford, L., Keith, T. Z., Floyd, R. G., Fields, C., & Schrank, F. A. (2003). Using the Woodcock-Johnson III Tests of Cognitive Abilities with students with attention deficit/hyperactivity disorder. In F. A. Schrank & D. P. Flanagan (Eds.), *WJ III Clinical Use and Interpretation: Scientist-Practitioner Perspectives* (pp. 320-342). Academic Press.

Ford-Jones, P. C. (2015). Misdiagnosis of attention deficit hyperactivity disorder: ‘Normal behaviour’and relative maturity. *Pediatrics & child health, 20*(4), 200-202. https://doi.org/10.1093/pch/20.4.200

Goldstein, S., & Naglieri, J. A. (2015). Using the CAS2 in the comprehensive assessment of ADHD. *The ADHD Report, 23*(7), 8-10. https://doi.org/10.1521/adhd.2015.23.7.8

Gray, J. A. (1994). Personality dimensions and emotion systems. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion* (pp.329-331). Oxford University Press.

Gyldrenkaerne, P., Dillion, H., Sharma, M., & Purdy, S. C. (2014). Attend to this: The relationship between auditory processing disorders and attention deficits. *Journal of the American Academy of Audiology. 25*(7), 676-687.

Hale, J. B., Reddy, L. A., Decker, S. L., Thompson, R., Henzel, J., Teodori, A., Forrest, E., Eusebio, E., & Denckla, M. B. (2009). Development and validation of an Attention-Deficit/ Hyperactivity Disorder (ADHD) executive function and behavior rating screening battery. *Journal of Clinical and Experimental Neuropsychology, 31*(8), 897-912. https://doi.org/10.1080/13803390802687423

Hale, J. B., Reddy, L. A., Wilcox, G., McLaughlin, A., Hain, L., Stern, A., Henzel, J., & Eusebio, E. (2010). Best practices in assessing and intervening with ADD/ADHD children and children with other frontal-striatal circuit disorders. In D. C. Miller (Ed.), *Best practices in school neuropsychology: Guidelines for effective practice, assessment and evidence-based interventions.* John Wiley & Sons.

Harrier, L., & DeOrnellas, K. (2005). Performance of children diagnosed with ADHD on selected planning and reconstitution tests. *Applied Neuropsychology, 12*(2)*,* 106-119. https://doi.org/10.1207/s15324826an1202_6

Hart, J. S. (2020). *A CHC theoretical approach to examining cognitive and Academic Deficits Among Students with ADHD Using a Three-Battery Approach* (3999) [Doctoral dissertation, University of Nevada]. UNVL Theses, Dissertation, Professional Papers and Capstones. https://digitalscholarship.unlv.edu/thesesdissertations/3999

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded.* Psychological Assessment Resources.

Huberty, C. J. (1984). Issues in the use of discriminant analysis. *Psychological Bulletin, 95*(1), 156-171. https://doi.org/10.1037/0033-2909.95.1.156

EveryCRSReport. (2007). Individuals with Disabilities Education Act of 2004, P.L. 108-446. EveryCRSReport. https://www.everycrsreport.com/reports/RL32716.html

Keith, T. Z., & Reynolds, M. R. (2012). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 758 –799). Guilford Press.

Keith, T. Z., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock–Johnson III tests of cognitive abilities. *Intelligence, 36*(6), 502–525. https://doi.org/10.1016/j.intell.2007.11.001

Kempton, S., Vance, A., Maruff, P., Luk, E., Costin, J., & Pantelis, C. (1999). Executive function and attention deficit hyperactivity disorder: Stimulant medication and better executive performance in children. *Psychological Medicine, 29*(3), 527-538. https://doi.org/10.1017/S0033291799008338

Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: Extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment, 5*(4), 395-399. https://doi.org/10.1037/1040-3590.5.4.395

Koonce, D. A. (2007). Attention Deficit Hyperactivity Disorder assessment practices by practicing school psychologists. *Journal of Psychoeducational Assessment, 25*(4), 319-333. https://doi.org/10.1177/0734282906298264

Landau, S., Milich, R., & Widiger, T. A. (1991). Predictive power methods may be more helpful in making a diagnosis than sensitivity and specificity. Journal of Child and Adolescent Psychopharmacology, *1*(5), 343-351. https://doi.org/10.1089/cap.1991.1.343

Lerner, J., & Yasutake, D. (2000). School strategies. In P. Accardo, T.A. Blondis, B.Y. Whitman, & M. Stein (Eds.), *Attention Deficit Disorders and Hyperactivity in children and adults* (2nd ed.) (pp. 461-476). Marcel Dekker, Inc.

Lezak, M. D. (1995). *Neuropsychological Assessment, Third Edition.* Oxford University Press.

Locke, S., McGrew, K. S., & Ford, L. (2011). A multiple group confirmatory factor analysis of the structural invariance of the Cattell–Horn– Carroll theory of cognitive abilities across matched Canadian and U.S. samples. *WMF Press Bulletin,* (1), 1–18.

McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment Scales for Children and Adolescents.* Edumetric and Clinical Science.

McGill, R. J. (2015). Incremental criterion validity of the WJ-III COG clinical clusters: Marginal predictive effects beyond the general factor. *Canadian Journal of School Psychology, 30*(1), 51-63. https://doi.org/10.1177/0829573514553926

McQuade, J. D., Tomb, M., Hoza, B., Waschbusch, D. A., Hurt, E. A., & Vaughn, A. J. (2011). Cognitive deficits and positively biased self-perceptions in children with ADHD. *Journal of abnormal child psychology*, *39*(2), 307-319. https://doi.org/10.1007/s10802-010-9453-7

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52*(3)*,* 194-215. https://doi.org/10.1037/h0048070

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial: Professional manual.* Psychological Assessment Resources.

Milich, R., Widiger, T. A., & Landau, D. (1987). Differential diagnosis of attention deficit and conduct disorders using conditional probabilities. *Journal of Consulting and Clinical Psychology, 55*(5), 762-767. https://doi.org/10.1037/0022-006X.55.5.762

Naglieri, J. A., & Das, J. P. (1997). *Cognitive Assessment System.* Riverside.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014). *Cognitive Assessment System* (2nd ed.). Pro-Ed.

Naglieri, J. A., Goldstein, S., Iseman, J. S., & Schwebach, A. (2003). Performance of children with attention deficit hyperactivity disorder and anxiety/depression on the WISC-III and Cognitive Assessment System. *Journal of Psychoeducational Assessment, 21*(1)*,* 32-42. https://doi.org/10.1177/073428290302100103

Naglieri, J. A., Salter, C. J., & Edwards, G. H. (2004). Assessment of children with attention and reading difficulties using the PASS theory and Cognitive Assessment System. *Journal of Psychoeducational Assessment, 22*(2)*,* 93-105. https://doi.org/10.1177/073428290402200201

Paolitto, A. W. (1999). Clinical validation of the Cognitive Assessment System for children with ADHD. *ADHD Report, 7*(1)*,* 1-5.

Penny, A. M., Waschbusch, D. A., Carrey, N., & Drabman, R. S. (2005). Applying a psychoeducational perspective to ADHD. *Journal of Attention Disorders, 8*(4), 208-220. https://doi.org/10.1177/1087054705278746

Pottinger, L. S. (2002). *Identifying AD/HD subtypes using the Cognitive Assessment System and the NEPSY* [Unpublished doctoral dissertation]. University of North Texas.

Reale, L., Bartoli, B., Cartabia, M., Zanetti, M., Costantino, M. A., Canevini, M. P., Termine C., & Bonati, M. (2017). Comorbidity prevalence and treatment outcome in children and adolescents with ADHD. *European child & adolescent psychiatry*, *26*, 1443-1457. https://doi.org/10.1007/s00787-017-1005-z

Reddy, L. A., Braunstein, D., & Dumont, R. (2008). Use of the Differential Ability Scales for children with ADHD. *School Psychology Quarterly, 23*(1), 139-148. https://psycnet.apa.org/doi/10.1037/1045-3830.23.1.139

Reddy, L. A., & Hale, J. B. (2007). Inattentiveness. In A. R. Eisen (Ed.), *Treating childhood behavioral and emotional problems: A step-by-step evidence-based approach.* (pp. 156-211). Guilford Publications, Inc.

Reddy, L. A., Hale, J. B., & Brodzinsky, L. K. (2011). Discriminant validity of the Behavior Rating Inventory of Executive Function Parent Form for children with Attention-Deficit/ Hyperactivity Disorder. *School Psychology Quarterly*, *26*(1), 45-55. https://doi.org/10.1037/a0022585

Reddy, L. A., Weissman, A., & Hale, J. B. (2013). *Neuropsychological assessment and intervention for emotional and behavior disordered youth: An integrated step-by-step evidence-based approach.* American Psychological Association Press.

Rowland, J. E. (2013). *Discriminative ability of CHC factor scores from the WJ III tests of cognitive abilities in children with ADHD* [Doctoral dissertation, Oklahoma State University]. SHAREOK.

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*(4), 532-545. https://doi.org/10.1037/0894-4105.19.4.532

Schrank, F. A., Flanagan, D. P., Woodcock, R. W., & Mascolo, J. T. (2002). *Essentials of WJ III cognitive abilities assessment.* Wiley.

Semrud-Clikeman, M., Filipek, P. A., Biederman, J., Steingard, R., Kennedy, D., Renshaw, P., & Bekken, K. (1994). Attention-deficit hyperactivity disorder: magnetic resonance imaging morphometric analysis of the corpus callosum. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*(6), 875-881. https://doi.org/10.1097/00004583-199407000-00014

Spenceley, L. M., Wood, W. L., & Lovett, B. J. (2020). Using the Woodcock-Johnson IV tests of cognitive abilities to detect feigned ADHD. *Applied Neuropsychology: Adult.* Advance Online Publication. 1-9. https://doi.org/10.1080/23279095.2020.1748631

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). Harper Collins.

Tannock, R. (1998). Attention deficit hyperactivity disorder: Advances in cognitive, neurobiological, and genetic research. *Journal of Child Psychology and Psychiatry, 39*(1), 65-99. https://doi.org/10.1111/1469-7610.00304

Taub, G. E., Keith, T. Z., Floyd, R. G., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly*, *23*(2), 187-198. https://doi/10.1037/1045-3830.23.2.187

Taub, G. E., & McGrew, K. S. (2004). A confirmatory factor analysis of Cattell–Horn–Carroll theory and cross-age invariance of the Woodcock– Johnson Tests of Cognitive Abilities III. *School Psychology Quarterly, 19*(1), 72– 87. https://doi.org/10.1521/scpq.19.1.72.29409

Vanderwood, M. L., McGrew, K. S., Flanagan, D. P., & Keith, T. Z. (2002). The contribution of general and specific cognitive abilities to reading achievement. *Learning and Individual Differences*, *13*(2), 159-188. https://doi.org/10.1016/S1041-6080(02)00077-8

Van Luit, J. E. H., Kroesbergen, E. H., & Naglieri, J. A. (2005). Utility of the PASS theory and Cognitive Assessment System for Dutch children with and without ADHD. *Journal of Learning Disabilities, 38*(5), 434-439. https://doi.org/10.1177/00222194050380050601

Weis, R. (2017). *Introduction to abnormal child and adolescent psychology.* Sage Publications.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001a). *Woodcock-Johnson III: Tests of Achievement.* Riverside Publishing

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001b). *Woodcock-Johnson III: Tests of Cognitive Abilities.* Riverside Publishing