Journal of Attention Disorders
http://jad.sagepub.com/

Construct Validity and Diagnostic Utility of the Cognitive Assessment System for ADHD

Gary L. Canivez and Allison R. Gaboury

Journal of Attention Disorders published online 11 June 2013
DOI: 10.1177/1087054713489021

The online version of this article can be found at:
http://jad.sagepub.com/content/early/2013/06/11/1087054713489021

Published by:
SAGE
http://www.sagepublications.com

Journal of Attention Disorders
XX(X) 1–11

DOI: 10.1177/1087054713489021
jad.sagepub.com
SAGE

Construct Validity and Diagnostic Utility of the Cognitive Assessment System for ADHD

Gary L. Canivez[1] and Allison R. Gaboury[2]

Abstract

Objective: The Das–Naglieri Cognitive Assessment System (CAS) is a test of cognitive abilities based on the Planning, Attention, Simultaneous, and Successive Theory (PASS). Studies of CAS performance by children with ADHD typically show lowest performance on Planning and deficits on Attention, but normal Simultaneous and Successive processing. Such distinct group differences studies support construct validity and are necessary, but not sufficient, for establishing diagnostic utility. **Method:** Students meeting *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*) criteria for ADHD (n = 20) and matched controls (n = 20) were assessed with the CAS to examine distinct group differences *and* diagnostic utility of CAS in correctly classifying the individuals from both groups. **Results:** CAS PASS scores were significantly related to ADHD characteristics demonstrating both distinct group differences (with medium to large effect sizes) *and* diagnostic utility (with medium effect sizes). **Conclusion:** Support was observed for CAS PASS score characteristics previously observed among students with ADHD, and this is the first study to report on the diagnostic utility of CAS PASS scores. Given the small sample, additional large-scale studies and cross-validation is needed. *(J. of Att. Dis. 2013; XX(X) 1-XX)*

Keywords

CAS, construct validity, diagnostic utility, diagnostic efficiency statistics, ROC curve analysis

ADHD is one of the most common disorders of childhood with prevalence estimates ranging from 3% to 7% according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*; American Psychiatric Association [APA], 2000) and the Centers for Disease Control and Prevention (2005). ADHD is a disorder that includes developmentally inappropriate impulsivity, inattention, and overactivity. Frazier, Youngstrom, Glutting, and Watkins (2007) reported on the significant impact ADHD can have on the academic and occupational achievement. School psychologists are often involved in assessments where attention problems, impulsivity, and overactivity are key features of children's learning and behavioral difficulties. Assessments for ADHD may include structured diagnostic interviews, teacher and parent report behavior rating scales, direct observations, neuropsychological tests, and cognitive tests. The American Academy of Child & Adolescent Psychiatry (2007) noted the use of structured diagnostic interviews and behavior rating scales as best practices in clinical assessment of ADHD.

Cognitive ability or intelligence tests and their resulting profiles have been recommended in assessment of ADHD (Kaufman, 1994; Prifitera & Dersh, 1993; Sattler, 2008; Schwean & McCrimmon, 2008) as constructs such as working memory, processing speed, and executive functioning are related to ADHD. Kaufman (1994) noted that low performance on Arithmetic, Coding, and Digit Span (Freedom From Distractibility subtests) on the Wechsler Intelligence Scale for Children–Revised (WISC-R; Wechsler, 1974) might be an indicator of ADHD due to distractibility. Another profile noted by Kaufman involved the Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 1991) Symbol Search, Coding, Arithmetic, and Digit Span subtests (SCAD profile). Gussin and Javorsky (1995), however, failed to find significant FD profile differences between ADHD and non-ADHD participants. Mayes, Calhoun, and Crowell (1998) reported lower scores on WISC-III Arithmetic and Digit Span subtests (Freedom From Distractibility) among students with ADHD, as did Anastopolous, Spisto, and Maher (1994), and also noted the SCAD pattern in the majority of ADHD cases. Mayes and Calhoun (2006) also reported 100% of students with ADHD showed their lowest score on the Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV; Wechsler, 2003a) Working Memory Index

[1]Eastern Illinois University, Charleston, USA
[2]Puyallup School District, WA, USA

Corresponding Author:
Gary L. Canivez, Eastern Illinois University, 600 Lincoln Avenue, Charleston, IL 61920, USA.
Email: glcanivez@eiu.edu

(WMI) *or* Processing Speed Index (PSI). It was also noted that lower WISC-IV PSI and WMI scores than Verbal Comprehension Index (VCI) and Perceptual Reasoning Index (PRI) scores were obtained by students with ADHD in WISC-IV standardization (Wechsler, 2003b). While such studies examining group differences (ADHD vs. non-ADHD) in subtest or index score performance suggest utility, studies that have directly tested the accuracy of individual classification with these scores have not been favorable. Watkins, Kush, and Glutting (1997a, 1997b) and Devena and Watkins (2012), for example, found that various WISC-III or WISC-IV score patterns or profiles did not provide acceptable levels of individual diagnostic utility for learning or emotional problems, or for ADHD, despite previous studies finding group differences.

The Das–Naglieri Cognitive Assessment System (CAS; Naglieri & Das, 1997a) is a test of cognitive abilities based on the Planning, Attention, Simultaneous, and Successive Theory (PASS; Das, Naglieri, & Kirby, 1994) which itself is linked to Luria's (1966a, 1966b, 1973) three functional units of the brain (Unit 1: Attention, Unit 2: Successive and Simultaneous, Unit 3: Planning). PASS theory (Das et al., 1994; Naglieri & Das, 1997b) proposes that children with ADHD would, as Barkley (2003, 2006) suggested, be more impulsive (and less reflective) in their cognitive processing, which in turn would negatively impact planning processing. Attention difficulties would be expected to negatively affect attention processing. Studies of CAS performance by children with ADHD typically show lowest performance on Planning with concurrent deficits on Attention but normal Simultaneous and Successive processing scores (Crawford, 2002; Naglieri & Das, 1997b, Naglieri, Goldstein, Iseman, & Schwebach, 2003; Naglieri, Salter, & Edwards, 2004; Paolitto, 1999; Pottinger, 2002; Van Luit, Kroesbergen, & Naglieri, 2005). Such group differences studies provide support for the construct validity of the CAS via distinct group differences; however, such support is inadequate for determining the utility of the CAS in individual diagnostic decision making (Mullins-Sweatt & Widiger, 2009). Distinct group differences are necessary but not sufficient. To date, there have been no diagnostic utility investigations of the CAS. Differences between groups, however, are a necessary but not sufficient condition for diagnostic utility and use of a test. Diagnostic utility requires investigation of indexes such as overall correct classification, sensitivity, specificity, positive predictive power (PPP), negative predictive power (NPP), false positive rate, and false negative rate (Kessel & Zimmerman, 1993; Landau, Milich, & Widiger, 1991; Meehl & Rosen, 1955; Milich, Widiger, & Landau, 1987). Sensitivity refers to the probability that a person with a disorder obtains a positive test finding and specificity refers to the probability that a person without a disorder obtains a negative test finding. PPP (value) refers to the probability that a person with a positive test finding also had the disorder while NPP (value) refers to the probability a person with a negative test finding did not have the disorder (Kessel & Zimmerman, 1993; Watkins, 2009). Landau et al. (1991) and Milich et al. (1987) have recommended the use of PPP and NPP as more meaningful indexes of diagnostic utility of a test.

Diagnostic utility of a test or set of test scores involves examining formulae that significantly differentiate diagnostic groups and examining classification accuracy when applied to the individuals. If a test score or set of test scores do not provide acceptable PPP and NPP (Landau et al. 1991; Milich et al., 1987) or results in high false positive or false negative results, then the test or set of test scores may not be useful for individual diagnostic purposes. Cross-validation with a different sample is also important. One such study of diagnostic utility related to ADHD (Canivez & Sprouls, 2005) found the Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993) to have an overall correct classification of 96% and very high diagnostic efficiency statistics (Sensitivity = .98, PPP = .94, Specificity = .95, NPP = .98).

One method for examining diagnostic utility is the utilization of discriminant function analysis (DFA) or logistic regression to create mathematical formulae that maximize discrimination between the two distinct groups. Then, individuals are reclassified based on the formulae and compared with the actual group membership from which they came. Kessel and Zimmerman (1993) standardized the calculation of the varying conditional probabilities that emerge from diagnostic utility studies due to problems they identified in the extant literature and manuscript review process. They also argued for presentation of *all* diagnostic efficiency statistics rather than focusing on sensitivity and specificity. As has been pointed out, diagnostic efficiency statistics (conditional probabilities) are dependent on a number of factors such as the base rate of the problem and the cut score used to determine the presence or absence of the disorder (Treat & Viken, 2012). A method that is independent of base rates and cut scores is receiver operating characteristic (ROC) curve analysis, which comes from signal detection theory (McFall & Treat, 1999; Swets, 1996; Swets, Dawes, & Monahan, 2000; Treat & Viken, 2012). By examining the true positive rate versus the false positive rate across all possible cut scores, a curve may be fit and the area under that curve (AUC) estimated to provide an indication of the accuracy of the measure or set of scores.

To date, no studies have been conducted on the diagnostic utility of the CAS in correctly identifying children with ADHD or ADHD symptoms with either diagnostic efficiency statistics or ROC methods. The present study examined the construct validity of the CAS by examining distinct group differences and then examined the

Table 1. Participant Demographic Characteristics.

	ADHD		RMC	
	n	%	*n*	%
Sex				
Male	16	80	14	70
Female	4	20	6	30
Grade				
Kindergarten	6	30	0	0
First	7	35	11	55
Second	7	35	9	45
Race/ethnicity				
Caucasian/White	14	70	17	85
Multiracial	3	15	3	15
Hispanic/Latino	1	5	0	0
No response	2	10	0	0
Special education				
No SDI	17	85	20	100
SDI	3	15	0	0

Note: RMC = random and matched control, SDI = specially designed instruction.

diagnostic utility of CAS in correctly differentiating individuals within a group of children with significant ADHD symptoms from those within a regular classroom control group.

Method

Participants

Informed parental consent was obtained for a total sample of 40 elementary school students from suburban Pierce County, Washington. Participants ranged from kindergarten to second grade. Groups consisted of 20 children meeting diagnostic criteria for ADHD and 20 children randomly selected and matched (to the extent possible) on key variables to form a control (random and matched control [RMC]) group. Matching ADHD and control group children was attempted for sex, age, race, and special education classification; however, only 11 students in the ADHD group were accurately matched with a random child from the same classroom. As no parents of potential control group children in kindergarten consented to their child's participation, five first-grade students and one second-grade student were used as closest age and grade matches for the kindergarten students. Table 1 presents demographic characteristics of the two groups. An independent *t*-test for differences between means showed that children in the ADHD group were significantly younger ($M = 6.60$, $SD = 1.14$) than those in the RMC group ($M = 7.45$, $SD = .51$), $t(38) = -3.04$, $p = .004$, $d = .96$, as a result of matching limitations. Because age-based standard scores on the CAS were used, age differences might be mitigated. Another sampling issue was the failed attempt to completely match on student sex as well as age and in the end there were two fewer male and two more females in the control group.

Instruments

ASCA. The ASCA (McDermott et al., 1993) is a teacher report, behavior rating scale designed for use with all noninstitutionalized youths ages 5 through 17 (Grades K-12). The ASCA measures six core syndromes (Attention Deficit Hyperactive [ADH], Solitary Aggressive–Provocative [SAP], Solitary Aggressive–Impulsive [SAI], Oppositional Defiant [OPD], Diffident [DIF], and Avoidant [AVO]) and two supplementary syndromes (Delinquent [DEL] and Lethargic–Hypoactive [LEH]). Core syndromes are combined to form two composite indexes: Overactivity (ADH, SAP, SAI, and OPD syndromes) and Underactivity (DIF and AVO syndromes). Raw scores are transformed to normalized *T* scores with uniform percentile associations and higher scores reflecting problematic behaviors and pathology.

Extensive evidence for ASCA score reliability and validity is presented in the ASCA manual (McDermott, 1994) and in the extant literature. Internal consistency estimates (Canivez, 2004, 2006; Canivez & Beran, 2009; Canivez & Bohan, 2006; Canivez & Sprouls, 2010; McDermott, 1993, 1994), short-term stability estimates (Canivez, Perry, & Weller, 2001; McDermott, 1993, 1994), and interrater agreement estimates (Canivez & Watkins, 2002; Canivez, Watkins, & Schaefer, 2002; McDermott, 1993, 1994; Watkins & Canivez, 1997) have supported various types of reliability for ASCA scores. Evidence of ASCA scores convergent validity (Canivez & Bordenkircher, 2002; Canivez & Rains, 2002; McDermott, 1993, 1994), divergent validity (Canivez & Bordenkircher, 2002; Canivez, Neitzel, & Martin, 2005; Canivez & Rains, 2002; McDermott, 1993, 1994), discriminative/discriminant validity (Canivez & Sprouls, 2005; McDermott, 1993, 1994; McDermott et al., 1995), and factorial validity and factorial validity generalization (Canivez, 2004, 2006; Canivez & Beran, 2009; Canivez & Bohan, 2006; Canivez & Sprouls, 2010; McDermott, 1993, 1994) have also been reported. In general, psychometric characteristics of the ASCA are acceptable and meet standards for both group and individual decision making (Canivez, 2001; Hills, 1981; Salvia & Ysseldyke, 2001).

Learning Behaviors Scale (LBS). The LBS (McDermott, Green, Francis, & Stott, 1999) is a teacher report questionnaire designed and found to measure student behaviors related to effective learning. It is composed of 29 positively and negatively worded items (behaviors) to reduce response sets and is rated on a 3-point scale (0 = *does not apply*, 1 = *sometimes applies*, 2 = *most often applies*) (McDermott,

1999). Of the 29 items, 25 are used to produce a total score and the four subscales include Competence/Motivation (CM), Attitude Toward Learning (AL), Attention/Persistence (AP), and Strategy/Flexibility (SF). Total and subscale raw scores are converted to *T* scores ($M = 50$, $SD = 10$) based on the nationally representative standardization sample of 1,500 students aged 5 to 17 years and higher scores reflect better developed learning behaviors. McDermott (1999) summarized supportive psychometric evidence for the LBS and additional support for the factor structure (Canivez & Beran, 2011; Canivez, Willenborg, & Kearney, 2006; Worrell, Vandiver, and Watkins, 2001) and incremental validity (Schaefer & McDermott, 1999; Yen, Konold, & McDermott, 2004) has also been reported.

Structured Diagnostic Interview for Parents (SDIP). The SDIP (Barkley & Murphy, 2006) is based on the diagnostic criteria published in the *DSM-IV-TR* (APA, 2000). Previous research has utilized either the Diagnostic Interview for Children and Adolescents (DICA; Reich, 2000) or the Diagnostic Interview Schedule for Children Version IV (DISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000), both of which are much lengthier than required for the present study. The SDIP was administered to a parent (typically mother) of a child referred for school behavioral problems to provide an assessment of reported symptoms regarding ADHD or rule out ADHD based on parent report and was the first step in determining inclusion into the ADHD group.

CAS. The CAS (Naglieri & Das, 1997a) is based on the PASS theory and has a nationally representative standardization sample ($N = 2{,}200$). A global, Full Scale score is provided in addition to four factor-based composite Scale scores (Planning, Attention, Successive, and Simultaneous) that represent the PASS theory of intelligence (Das et al., 1994). PASS Scale scores and Full Scale score are all standard scores ($M = 100$, $SD = 15$) and the Standard Battery includes all 12 subtests (Naglieri & Das, 1997b). The Full Scale score internal consistency coefficients by age ranged from .95 to .97 and average internal consistency coefficients by age for the PASS Scales scores are .88 (Planning), .88 (Attention), .93 (Simultaneous), and .93 (Successive) (Naglieri & Das, 1997b). Additional psychometric information regarding various estimates of score validity is provided in the *Interpretive Handbook* (Naglieri & Das, 1997b). The CAS was recommended for use in diagnosing learning strengths and weaknesses, classification (learning disabilities, attention deficit disorder, mental retardation, giftedness), eligibility decisions (meeting state or federal criteria), and consideration of the appropriateness of treatment, instruction, or remedial programs (Naglieri & Das, 1997b) and was reported to be useful for testing special populations (Thompson, 2001).

Procedure

Permission to complete research was granted by the participating school district and university institutional review board (IRB) approval was also secured. A total of 78 (K-2) student referrals for behavior difficulties over a 2-year period were used to obtain the final ADHD sample. Following classroom teacher referral, a semistructured teacher interview concentrating on significant problem behaviors, behavior onset and duration, and behavioral intensity was conducted. There were 26 students excluded due to either not demonstrating ADHD symptoms, presenting ADHD *and* multiple other problem behaviors, or the school participation withdrawal.

Parents of students who reportedly demonstrated behaviors related to ADHD ($n = 52$) based on teacher referral and subsequent teacher interview were contacted to obtain informed consent for participation and for administration of the SDIP (Barkley & Murphy, 2006). Students ($n = 19$) whose parents declined completion of the SDIP or later withdrew participation were also excluded. The remaining 33 students met the inclusion criteria from the SDIP by meeting the minimum of six inattention and/or six hyperactive-impulsive symptoms. The majority of symptoms were required to be manifested under the hyperactive-impulsive type, to be inappropriate for the child's age, to have lasted at least the past 6 months, and to have caused some impairment prior to 7 years of age. Furthermore, symptoms were presently causing impairment in home and school (a cause of referral) and produced evidence of clinically significant impairment in social or academic functioning. Symptoms also were not reported to be occurring only during a Pervasive Developmental Disorder or Psychotic Disorder or better accounted for by another mental disorder.

Teachers of these 33 students then completed the ASCA and students who obtained an ADH scale *T* score > 65 ($n = 20$) composed the ADHD group. The ASCA was selected as an additional criterion measure due to its diagnostic utility demonstrated in a previous study (Canivez & Sprouls, 2005). Teachers also completed the LBS (McDermott et al., 1999) to measure behaviors related to effective learning processes for descriptive purposes. Teachers were not provided SDIP results and were blind to diagnostic group inclusion prior to completing both the ASCA and LBS.

A RMC group ($n = 20$) was created with children from the same classroom as the student in the ADHD group to the extent possible given the voluntary nature of the study. Unfortunately, none of the parents of kindergarten children who were potential control group participants agreed to allow their children to participate in the study. Thus, it was necessary to use somewhat older students in Grade 1 ($n = 5$) and Grade 2 ($n = 1$) as matches to the kindergarten children in the ADHD group. Teachers completed ASCA and LBS for students in the RMC group but parents were

Table 2. MANOVA and ANOVA Results for ASCA Core Syndromes, LBS Factors, and CAS PASS Scales.

	SS	SS Error	MS	MS Error	*F*	*p*	η^2
ASCA							
ADH	7,128.90	1,646.20	7,128.90	43.32	164.56	.001	.812
SAP	3,900.63	2,744.15	3,900.63	72.21	54.01	.001	.587
SAI	2,924.10	2,667.80	2,924.10	70.21	41.65	.001	.523
OPD	2,640.63	3,419.75	2,640.63	89.99	29.34	.001	.436
DIF	10.00	2,441.60	10.00	64.25	0.16	.695	.004
AVO	156.03	3,617.75	156.03	95.20	0.21	.208	.041
LBS							
CM	180.63	5,428.35	180.63	142.85	1.26	.268	.032
AL	1,276.90	4,462.60	1,276.90	117.44	10.87	.002	.222
AP	4,182.03	5,727.95	4,182.03	150.74	27.74	.001	.422
SF	6,528.03	5,309.35	6,528.03	139.72	46.72	.001	.551
CAS							
Planning	1,562.50	10,037.90	1,562.50	264.16	5.92	.020	.135
Attention	2,907.03	7,135.95	2,907.03	187.79	15.48	.001	.289
Simultaneous	819.03	10,189.95	819.03	268.16	3.05	.089	.074
Successive	32.40	6,356.00	32.40	167.26	0.19	.662	.005

Note: ASCA = Adjustment Scales for Children and Adolescents; CAS = Cognitive Assessment System; LBS = Learning Behaviors Scale; PASS = Planning, Attention, Simultaneous, and Successive; ADH = Attention Deficit Hyperactive; SAP = Solitary Aggressive–Provocative; SAI = Solitary Aggressive–Impulsive; OPD = Oppositional Defiant; DIF = Diffident; AVO = Avoidant; CM = Competence/Motivation; AL = Attitude Toward Learning; AP = Attention/Persistence; SF = Strategy/Flexibility. MANOVA for ASCA Core Syndromes: Wilks's λ = .17, $F(6, 33) = 27.92$, $p = .0001$, multivariate effect size = .835, power = 1.0. ASCA ANOVA *df* (1, 38). MANOVA for LBS factors: Wilks's λ = .38, $F(4, 35) = 14.61$, $p = .0001$, multivariate effect size = .625, power = 1.0. LBS ANOVA *df* (1, 38). MANOVA for CAS PASS Scales: Wilks's λ = .70, $F(4, 35) = 3.83$, $p = .011$, multivariate effect size = .304, power = .849. CAS ANOVA *df* (1, 38).

not interviewed using the SDIP. Finally, students from both groups were individually administered the CAS Standard Battery.

Analyses

MANOVA and ANOVA for distinct group differences were used to analyze mean differences on the ASCA, LBS, and CAS PASS scale scores. An independent *t*-test for differences between means was included to determine group differences on the CAS Full Scale score. To examine the diagnostic utility of the CAS, direct DFA (Tabachnick & Fidell, 2007) with subsequent diagnostic efficiency statistics (Canivez & Watkins, 1996; Kessel & Zimmerman, 1993) were used. Diagnostic efficiency statistics estimated sensitivity (probability of a person with disorder having a positive test score), specificity (probability of a person without disorder having a negative test score), PPP (probability of a person with a positive test score having the disorder) and NPP (probability of a person with a negative test score not having the disorder), false positive and negative rates, and overall correct classification (Canivez & Watkins, 1996; Watkins, 2009) as recommended by Kessel and Zimmerman (1993). In addition, ROC curve analysis was conducted using the ROC program provided by Watkins (2002) to estimate AUC for further examination of diagnostic utility because diagnostic efficiency statistics are affected by cut scores and base rates as well as costs and benefits (Treat & Viken, 2012). MANOVA, ANOVA, and DFA were conducted with SPSS version 17.0 for Macintosh while diagnostic efficiency statistics were calculated using the Automated Calculation of Diagnostic Efficiency Statistics spreadsheet (Canivez, 1994). Because bias enters into classifications when the individual cases are used to both create discriminant classification coefficients *and* then be statistically assigned to groups based on those coefficients, jackknifed classification (leave one out) was used to provide a method of cross-validation for comparison with the discriminant function and diagnostic efficiency statistics results (Tabachnick & Fidell, 2007).

Results

Distinct Group Differences

ASCA. Table 2 presents results of the one-way MANOVA and subsequent univariate ANOVAs for differences between the ADHD and RMC groups on the ASCA core syndromes. Statistically significant group differences with large effect sizes were observed for the ADH, SAP, SAI, and OPD syndromes and no statistically significant differences were observed on the DIF and AVO syndromes. Table 3 presents descriptive statistics illustrating group

Table 3. Descriptive Statistics, *p*, and Effect Size Estimates for ASCA Core Syndromes, LBS Factors, and CAS PASS Score Differences Between the ADHD and RMC Groups.

	ADHD		RMC			
	M	*SD*	*M*	*SD*	*p*	*d*
ASCA						
ADH	72.00	4.37	45.30	8.22	.001	4.06
SAP	66.95	9.93	47.20	6.77	.001	2.32
SAI	64.10	11.85	47.00	0.00	.001	2.04
OPD	61.25	11.85	45.00	6.29	.001	1.71
DIF	45.10	7.18	46.10	8.77	.695	0.13
AVO	50.55	9.92	46.60	9.59	.208	0.41
LBS						
CM	43.90	7.22	48.15	15.28	.268	0.36
AL	40.60	5.25	51.90	14.40	.002	1.04
AP	30.05	13.98	50.50	10.30	.001	1.67
SF	27.60	14.20	53.15	8.83	.001	2.16
CAS						
Planning	91.55	16.57	104.05	15.93	.020	0.77
Attention	90.45	12.81	107.50	14.54	.001	1.24
Simultaneous	102.50	17.55	111.55	15.11	.089	0.55
Successive	97.30	13.69	99.10	12.14	.662	0.14

Note: ASCA = Adjustment Scales for Children and Adolescents; LBS = Learning Behaviors Scale; CAS = Cognitive Assessment System; PASS = PASS = Planning, Attention, Simultaneous, and Successive; RMC = random and matched control; ADH = Attention Deficit Hyperactive; SAP = Solitary Aggressive–Provocative; SAI = Solitary Aggressive–Impulsive; OPD = Oppositional Defiant; DIF = Diffident; AVO = Avoidant; CM = Competence/Motivation; AL = Attitude Toward Learning; AP = Attention/Persistence; SF = Strategy/Flexibility. *d* = Cohen's *d* effect size estimate (Cohen, 1988).

differences on the ASCA and the largest difference between the ADHD and RMC groups was on the ADH syndrome with the mean score for the ADHD group in the maladaptive range (McDermott, 1993). Also, mean scores for the ADHD group on the SAP, SAI, and OPD syndromes were in the at-risk range (McDermott, 1993). Mean DIF and AVO scores for the ADHD group were in the adjusted range (McDermott, 1993). All mean ASCA scores for the RMC group were in the adjusted (average) range.

LBS. Table 2 presents results of the one-way MANOVA and subsequent univariate ANOVAs for differences between the ADHD and RMC groups on the LBS scales. Statistically significant group differences with large effect sizes were observed for the AL, AP, and SF scales but no significant difference was observed on the CM scale. Table 3 presents descriptive statistics illustrating group differences on the LBS and the largest difference was on the SF scale where the ADHD group mean score was more than 2 *SD* below average. The ADHD group also showed mean ratings on the AP scale nearly 2 *SD* below average. Mean LBS ratings for the RMC group were all within the average range.

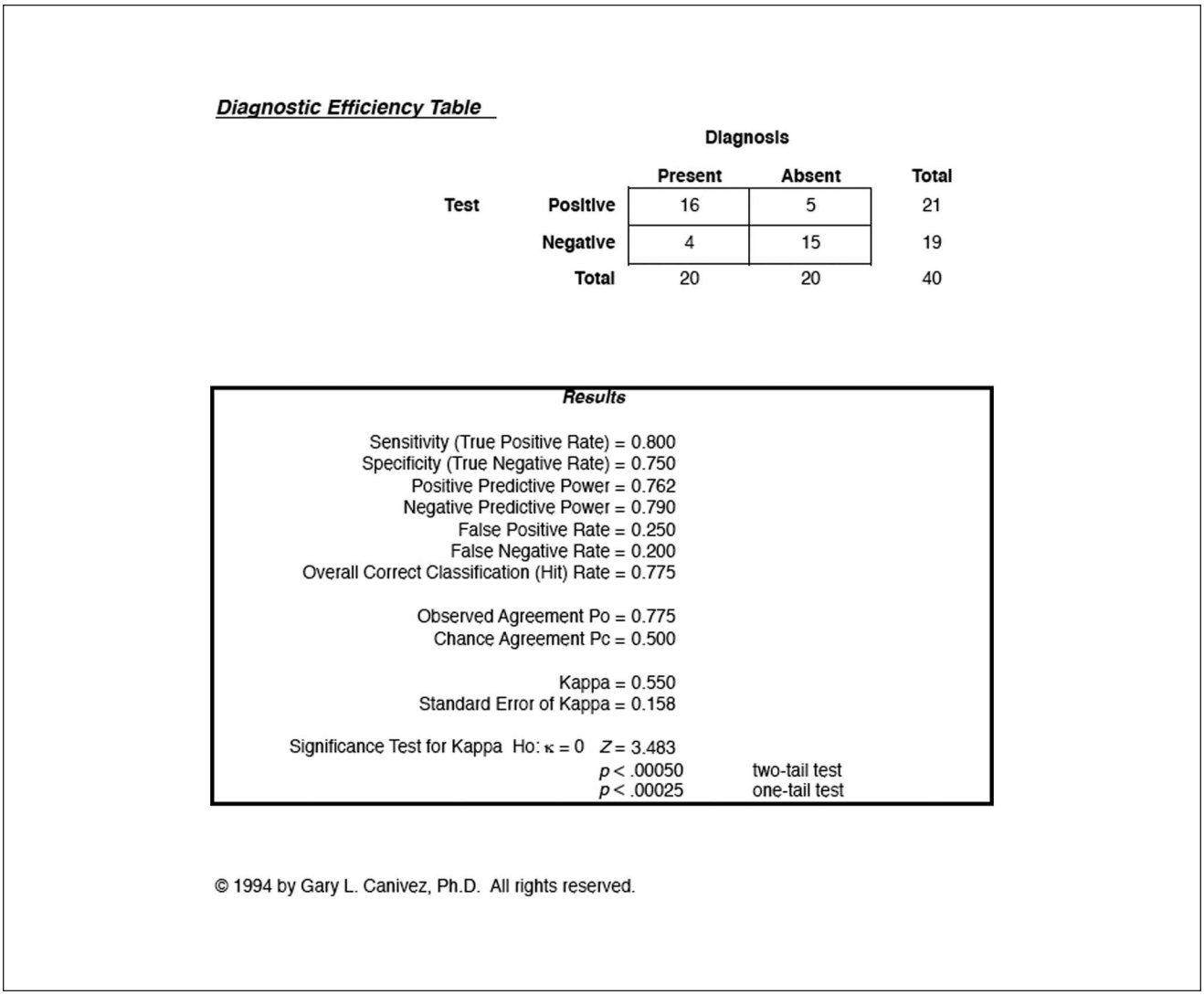
CAS. Table 2 presents results of the one-way MANOVA and subsequent univariate ANOVAs for differences between the ADHD and RMC group on the CAS PASS scales. MANOVA produced a statistically significant difference between the ADHD and RMC groups on the PASS scores overall. Subsequent one-way univariate ANOVAs produced statistically significant differences between the ADHD and RMC groups on the Planning and Attention scales but *not* on the Simultaneous or Successive scales. Table 3 presents descriptive statistics illustrating group differences on the CAS. The ADHD group obtained significantly lower scores than RMC group on the CAS Planning (medium effect size) and Attention (large effect size) scales. Table 3 presents descriptive statistics illustrating group differences on the CAS. The independent *t*-test for differences between means showed that the ADHD group (M = 93.75, SD = 17.17) had significantly lower CAS Full Scale scores than the RMC group (M = 107.05, SD = 14.52), $t(38) = -2.65$, $p = .012$, $d = .84$ (large effect; Cohen, 1988); a result of significantly lower scores on the Planning and Attention scales.

Discriminative Validity/Diagnostic Utility

Direct DFA was statistically significant; Wilks's $\lambda = .696$, $\chi^2(4) = 13.07$, $p = .011$. Diagnostic efficiency statistics (see Figure 1) based on comparing DFA reclassification of individuals in the ADHD and RMC groups with the original a priori classification produced overall correct classification of 77.5%. Diagnostic efficiency statistics were supportive (Sensitivity = .80, PPP = .76, Specificity = .75, NPP = .79, False Positive Rate = .25, False Negative Rate = .20). Jackknifed classification (leave one out) was used to cross-validate DFA comparisons and diagnostic efficiency statistics are presented in Figure 2. Overall correct classification for the jackknifed classification method was 70% and illustrated some shrinkage typically observed in cross-validation. All diagnostic efficiency statistics were reduced in cross-validation. ROC curve analysis based on the DFA discriminant scores (see Figure 3) produced medium diagnostic accuracy (Swets, 1988) with an AUC = .846 (SE_{AUC} = .063; 95% confidence interval $[CI]_{AUC}$ = [.722, .970]). This figure indicates that about 85% of the time, a randomly selected child from this ADHD sample would have a lower discriminant score (lower CAS performance) than a randomly selected child from the RMC sample.

Discussion

This study is the first to examine the diagnostic utility of CAS PASS scores, thereby going beyond the typical distinct group differences or discriminant validity studies. Results suggested that the CAS may be a useful cognitive measure for assisting in diagnosing ADHD in young children and results were consistent with CAS and ADHD theory. Consistent with previous research (Crawford, 2002;



Diagnostic Efficiency Table

		Diagnosis: Present	Diagnosis: Absent	Total
Test	Positive	16	5	21
	Negative	4	15	19
	Total	20	20	40

Results

Sensitivity (True Positive Rate) = 0.800
Specificity (True Negative Rate) = 0.750
Positive Predictive Power = 0.762
Negative Predictive Power = 0.790
False Positive Rate = 0.250
False Negative Rate = 0.200
Overall Correct Classification (Hit) Rate = 0.775

Observed Agreement Po = 0.775
Chance Agreement Pc = 0.500

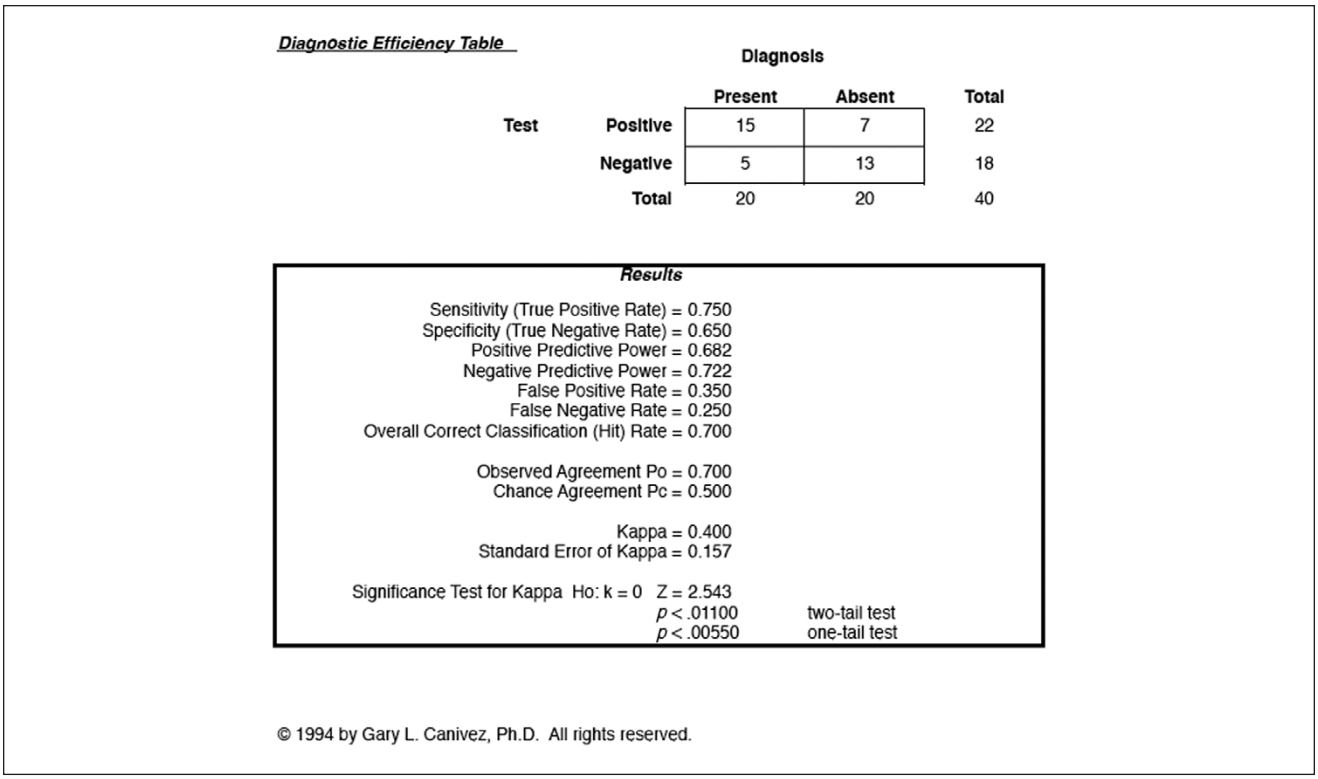
Kappa = 0.550
Standard Error of Kappa = 0.158

Significance Test for Kappa Ho: $\kappa = 0$ $Z = 3.483$
$p < .00050$ two-tail test
$p < .00025$ one-tail test

Figure 1. Diagnostic efficiency table comparing CAS direct DFA classifications (test) to diagnostic groups formed a priori based on SDIP and ASCA ADH criteria (diagnosis).
Note: CAS = Cognitive Assessment System; DFA = discriminant function analysis; SDIP = Structured Diagnostic Interview for Parents; ASCA = Adjustment Scales for Children and Adolescents; ADH = Attention Deficit Hyperactive.

Naglieri & Das, 1997b, Naglieri et al., 2003; Naglieri et al., 2004; Paolitto, 1999; Pottinger, 2002; Van Luit et al., 2005), the present study found students in the ADHD group to obtain significantly lower CAS performance on the Planning and Attention scales and medium to large effect sizes, but unlike previous studies, the present sample obtained somewhat lower scores on the Attention scale than Planning. This result may be sample specific and given the small sample could reflect sampling error.

Whereas previous studies have not investigated the ability of the CAS to correctly identify individual members of diagnostic groups, the present study demonstrated the potential of the CAS to correctly identify individual students who demonstrated behaviors consistent with ADHD symptoms and problems measured in several ways (semistructured and structured interviews and teacher report behavior rating scale). While shrinkage was observed when jackknifed classification was used in cross-validation, diagnostic efficiency statistics were still statistically significant and generally good. The PPP of 76.2%, the more important diagnostic efficiency statistic for inclusion (Landau et al., 1991), was high as suggested by Landau et al. (1991) and ROC analyses also showed a moderate degree of diagnostic accuracy and is not affected by base rates (50% in the present study) or cut scores. The diagnostic efficiency statistics (PPP and NPP) in the present study were as high or higher than those reported by Doyle, Biederman, Seidman, Weber, and Faraone (2000) in examining the diagnostic utility of a battery of neuropsychological



Diagnostic Efficiency Table

		Diagnosis: Present	Diagnosis: Absent	Total
Test	Positive	15	7	22
	Negative	5	13	18
	Total	20	20	40

Results

Sensitivity (True Positive Rate) = 0.750
Specificity (True Negative Rate) = 0.650
Positive Predictive Power = 0.682
Negative Predictive Power = 0.722
False Positive Rate = 0.350
False Negative Rate = 0.250
Overall Correct Classification (Hit) Rate = 0.700

Observed Agreement Po = 0.700
Chance Agreement Pc = 0.500

Kappa = 0.400
Standard Error of Kappa = 0.157

Significance Test for Kappa Ho: k = 0 Z = 2.543
$p < .01100$ two-tail test
$p < .00550$ one-tail test

Figure 2. Diagnostic efficiency table comparing "jackknifed" (leave one out) CAS direct DFA classifications (test) to diagnostic groups formed a priori based on SDIP and ASCA ADH criteria (diagnosis).
Note: CAS = Cognitive Assessment System; DFA = discriminant function analysis; SDIP = Structured Diagnostic Interview for Parents; ASCA = Adjustment Scales for Children and Adolescents; ADH = Attention Deficit Hyperactive.

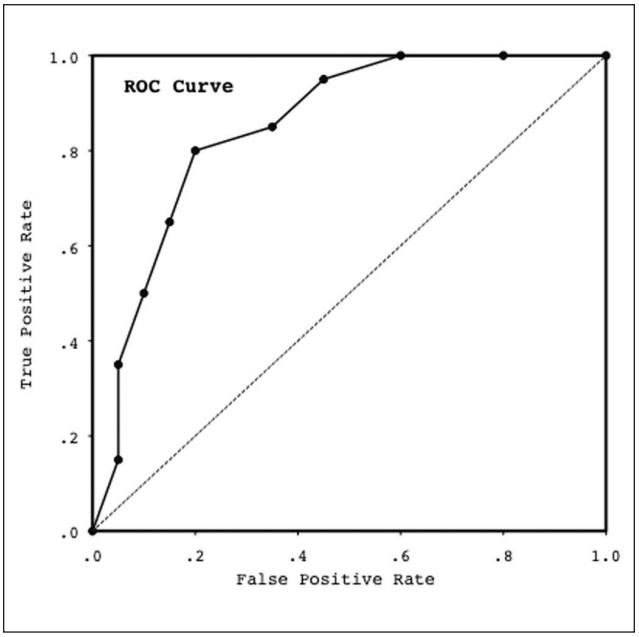


Figure 3. CAS diagnostic utility ROC curve.
Note: CAS = Cognitive Assessment System; ROC = receiver operating characteristic; AUC = area under the curve. AUC = .846, SE_{AUC} = .063, asymptotic 95% confidence interval = [.722, .970], $p < .001$.

tests in identifying ADHD, and AUC estimates for the CAS in the present study were higher than those in Doyle et al.

Conclusions and generalization of findings must be tempered by the significant limitations of the extremely small and demographically restricted sample from the Pacific Northwest as well as the fact that control group participants for kindergarten children from the ADHD group *were not* from the same classroom or from other kindergarten classrooms. Matching for these youngsters was as best as could be obtained given the site restrictions and voluntary nature of the procedure but effects might be mitigated by the use of age-based CAS standard scores. Another consideration is the fact that the ASCA produced higher diagnostic efficiency statistics in discriminating ADHD in a larger sample (Canivez & Sprouls, 2005). While not included in the Canivez and Sprouls (2005) study, ROC analysis based on discriminant scores was subsequently conducted to determine AUC produced by the ASCA and AUC = .985 (SE_{AUC} = .012). Also, discrimination in the present study was based on a base rate of 50% due to matched samples, so examination of CAS discrimination when the ADHD sample base rate is closer to the population base rate (3%-7%) is needed.

Future studies should continue to investigate the diagnostic utility of CAS and with much larger samples of children

referred for evaluation of ADHD. Furthermore, it would be useful to examine the utility of CAS in differential diagnosis when attempting to differentiate children with ADHD from other disruptive behavior disorders such as oppositional defiant disorder (ODD) and conduct disorder (CD); a more difficult but important test of utility. Should the CAS be capable of differentiating ADHD from other externalizing disorders such as ODD and CD at levels comparable with this study then CAS application in an actuarial classification might be advocated (Meehl, 1956; Meehl & Rosen, 1955). Other important investigations could be to assess the extent that cognitive measures such as the CAS provide incremental predictive validity compared with structured diagnostic interviews and behavior rating scales or incremental diagnostic utility when added to diagnostic interviews and behavior rating scales such as the ASCA, which was shown to possess higher diagnostic utility in ADHD (Canivez & Sprouls, 2005). It is hoped that such examinations will be conducted with the soon to be published revision of the CAS.

Authors' Note

This article is based, in part, on Allison R. Gaboury's specialist in school psychology thesis.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

American Academy of Child & Adolescent Psychiatry. (2007). Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 894-921. doi:10.1097/chi.0b013e318054e724

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Anastopolous, 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. doi:10.1037/1040-3590.6.4.368

Barkley, R. A. (2003). Attention-deficit/hyperactivity disorder. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (2nd ed., pp. 75-143). New York, NY: Guilford.

Barkley, R. A. (2006). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3rd ed.). New York, NY: Guilford.

Barkley, R. A., & Murphy, K. R. (2006). *Attention-deficit hyperactivity disorder: A clinical workbook* (3rd ed.). New York, NY: Guilford.

Canivez, G. L. (1994). *Automated calculation of diagnostic efficiency statistics* (Microsoft® Excel™ Spreadsheet Template for the Apple® Macintosh™ Microcomputer). Phoenix, AZ: Author.

Canivez, G. L. (2001). Review of the Adjustment Scales for Children and Adolescents. In J. C. Impara & B. Plake (Eds.), *The fourteenth mental measurement yearbook* (pp. 22-24). Lincoln: Buros Institute of Mental Measurements, University of Nebraska.

Canivez, G. L. (2004). Replication of the Adjustment Scales for Children and Adolescents core syndrome factor structure. *Psychology in the Schools*, *41*, 191-199. doi:10.1002/pits.10121

Canivez, G. L. (2006). Adjustment Scales for Children and Adolescents and native American Indians: Factorial validity generalization for Ojibwe youths. *Psychology in the Schools*, *43*, 685-694. doi:10.1002/pits.20179

Canivez, G. L., & Beran, T. N. (2009). Adjustment Scales for Children and Adolescents: Factorial validity in a Canadian sample. *Canadian Journal of School Psychology*, *24*, 284-302. doi:10.1177/0829573509344344

Canivez, G. L., & Beran, T. N. (2011). Learning Behaviors Scale and Canadian youths: Factorial validity generalization and comparisons to the U.S. standardization sample. *Canadian Journal of School Psychology*, *26*, 193-208. doi:10.1177/0829573511416440

Canivez, G. L., & Bohan, K. (2006). Adjustment Scales for Children and Adolescents and native American Indians: Factorial validity generalization for Yavapai Apache Youths. *Journal of Psychoeducational Assessment*, *24*, 329-341. doi:10.1177/0734282906291397

Canivez, G. L., & Bordenkircher, S. E. (2002). Convergent and divergent validity of the Adjustment Scales for Children and Adolescents for Children and the Preschool and Kindergarten Behavior Scales. *Journal of Psychoeducational Assessment*, *20*, 30-45. doi:10.1177/073428290202000103

Canivez, G. L., Neitzel, R., & Martin, B. E. (2005). Construct validity of the Kaufman Brief Intelligence Test, Wechsler Intelligence Scale for Children–Third Edition, and Adjustment Scales for Children and Adolescents. *Journal of Psychoeducational Assessment*, *23*, 15-34. doi:10.1177/073428290502300102

Canivez, G. L., Perry, A., & Weller, E. (2001). Stability of the Adjustment Scales for Children and Adolescents. *Psychology in the Schools*, *38*, 217-227. doi:10.1002/pits.1012

Canivez, G. L., & Rains, J. D. (2002). Construct validity of the Adjustment Scales for Children and Adolescents and the Preschool and Kindergarten Behavior Scales: Convergent and divergent evidence. *Psychology in the Schools*, *39*, 621-633. doi:10.1002/pits.10063

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

Canivez, G. L., & Sprouls, K. (2010). Adjustment Scales for Children and Adolescents: Factorial validity generalization with Hispanic/Latino youths. *Journal of Psychoeducational Assessment*, *28*, 209-221. doi:10.1177/0734282909349213

Canivez, G. L., & Watkins, M. W. (1996). Automated calculation of diagnostic efficiency statistics. *Behavior Research*

Methods, Instruments, & Computers, *28*, 132-133. doi:10.3758/BF03203650

Canivez, G. L., & Watkins, M. W. (2002). Interrater agreement for syndromic profile classifications on the Adjustment Scales for Children and Adolescents. *Assessment for Effective Intervention*, *28*, 39-46. doi:10.1177/073724770202800105

Canivez, G. L., Watkins, M. W., & Schaefer, B. A. (2002). Interrater agreement for discriminant classifications for the Adjustment Scales for Children and Adolescents. *Psychology in the Schools*, *39*, 375-384. doi:10.1002/pits.10043

Canivez, G. L., Willenborg, E., & Kearney, A. (2006). Replication of the Learning Behaviors Scale factor structure with an independent sample. *Journal of Psychoeducational Assessment*, *24*, 97-111. doi:10.1177/0734282905285239

Centers for Disease Control and Prevention. (2005). Mental health in the United States: Prevalence of diagnosis and medication treatment for attention-deficit hyperactivity disorder—United States, 2003. *Morbidity and Mortality Weekly Report*, *54*, 842-847.

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

Crawford, E. N. (2002). Profiles for exceptional samples on the Cognitive Assessment System using configural frequency analysis. *Dissertation Abstracts International: Section B: Sciences and Engineering*, *63*, 3061.

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

Devena, S. E., & Watkins, M. W. (2012). Diagnostic utility of WISC-IV General Abilities Index and Cognitive Proficiency Index difference scores among children with ADHD. *Journal of Applied School Psychology*, *28*, 133-154. doi:10.1080/15377903.2012.669743

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*, 477-488. doi:10.1037/0022-006X.68.3.477

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and achievement: Meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. *Journal of Learning Disabilities*, *40*, 49-65. doi:10.1177/00222194070400010401

Gussin, B., & Javorsky, J. (1995). The utility of the WISC-III freedom from distractibility in the diagnosis of youth with attention deficit hyperactivity disorder in a psychiatric sample. *Diagnostique*, *21*, 29-40.

Hills, J. R. (1981). *Measurement and evaluation in the classroom* (2nd ed.). Columbus, OH: Merrill.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York, NY: Wiley.

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*, 395-399. doi:10.1037/1040-3590.5.4.395

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*, 343-351. doi:10.1089/cap.1991.1.343

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

Luria, A. R. (1966b). *Human brain and psychological processes*. New York, NY: Harper & Row.

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

Mayes, S. D., & Calhoun, S. L. (2006). WISC-IV and WISC-III profiles in children with ADHD. *Journal of Attention Disorders*, *9*, 486-493. doi:10.1177/1087054705283616

Mayes, S. D., Calhoun, S. L., & Crowell, E. W. (1998). WISC-III freedom from distractibility as a measure of attention in children with and without attention deficit hyperactivity disorder. *Journal of Attention Disorders*, *2*, 217-227. doi:10.1177/108705479800200402

McDermott, P. A. (1993). National standardization of uniform multisituational measures of child and adolescent behavior pathology. *Psychological Assessment*, *5*, 413-424. doi:10.1037/1040-3590.5.4.413

McDermott, P. A. (1994). *National profiles in youth psychopathology: Manual of Adjustment Scales for Children and Adolescents*. Philadelphia, PA: Edumetric and Clinical Science.

McDermott, P. A. (1999). National scales of differential learning behaviors among American children and adolescents. *School Psychology Review*, *28*, 280-291.

McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (1999). *Learning Behaviors Scale*. Philadelphia, PA: Edumetric and Clinical Science.

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

McDermott, P. A., Watkins, M. W., Sichel, A. F., Weber, E. M., Keenan, J. T., Holland, A. M., & Leigh, N. M. (1995). The accuracy of new national scales for detecting emotional disturbance in children and adolescents. *Journal of Special Education*, *29*, 337-354. doi:10.1177/002246699502900306

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology*, *50*, 215-241. doi:10.1146/annurev.psych.50.1.215

Meehl, P. E. (1956). Wanted—A good cookbook. *American Psychologist*, *11*, 263-272. doi:10.1037/h0044164

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194-216. doi:10.1037/h0048070

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*, 762-767. doi:10.1037/0022-006X.55.5.762

Mullins-Sweatt, S. N., & Widiger, T. A. (2009). Clinical utility and DSM-V. *Psychological Assessment*, *21*, 302-312. doi:10.1037/a0016607

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

Naglieri, J. A., & Das, J. P. (1997b). *Cognitive Assessment System: Interpretive handbook*. Itasca, IL: Riverside.

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*, 32-42. doi: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*, 93-105. doi:10.1177/073428290402200201

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

Pottinger, L. S. (2002). Identifying AD/HD subtypes using the Cognitive Assessment System and the NEPSY. *Dissertation Abstracts International: Section B: Sciences and Engineering*, *63*, 1012.

Prifitera, A., & Dersh, J. (1993). Base rates of WISC-III diagnostic subtest patterns among normal, learning-disabled, and ADHD samples. *Journal of Psychoeducational Assessment, WISC-III Monograph*, *11*, 43-55.

Reich, W. (2000). Diagnostic interview for children and adolescents (DICA). *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 59-66. doi:10.1097/00004583-200001000-00017

Salvia, J., & Ysseldyke, J. E. (2001). *Assessment* (8th ed.). Boston, MA: Houghton Mifflin.

Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego, CA: Author.

Schaefer, B. A., & McDermott, P. A. (1999). Learning behavior and intelligence as explanations for children's scholastic achievement. *Journal of School Psychology*, *37*, 299-313. doi:10.1016/S0022-4405(99)00007-2

Schwean, V. L., & McCrimmon, A. (2008). Attention-deficit/hyperactivity disorder: Using the WISC-IV to inform intervention planning. In A. Prifitera, D. H. Saklofske & L. G. Weiss (Eds.), *WISC-IV clinical assessment and intervention* (pp. 193-215). San Diego, CA: Academic Press.

Shaffer, D. F., 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. doi:10.1097/00004583-200001000-00014

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*, 1285-1293. doi:10.1126/science.3287615

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychological diagnostics: Collected papers*. Mahwah, NJ: Lawrence Erlbaum.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*, 1-26. doi:10.1111/1529-1006.001

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Thompson, D. (2001). Review of the Das–Naglieri Cognitive Assessment System. In J. C. Impara & B. Plake (Eds.), *The fourteenth mental measurement yearbook* (pp. 368-370). Lincoln: Buros Institute of Mental Measurements, University of Nebraska.

Treat, T. A., & Viken, R. J. (2012). Measuring test performance with signal detection theory techniques. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf & K. J. Sher (Eds.), *Handbook of research methods in psychology, Vol. 1: Foundations, planning, measures, and psychometrics* (pp. 723-744). Washington, DC: American Psychological Association.

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*, 434-439. doi:10.1177/00222194050380050601

Watkins, M. W. (2002). *ROC analysis* [Computer software]. State College, PA: Ed & Psych Associates.

Watkins, M. W. (2009). Errors in diagnostic decision making and clinical judgment. In T. B. Gutkin & C. R. Reynolds (Eds.), *Handbook of school psychology* (pp. 210-229). Hoboken, NJ: John Wiley.

Watkins, M. W., & Canivez, G. L. (1997). Interrater agreement of the Adjustment Scales for Children and Adolescents. *Diagnostique*, *22*, 205-213.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997a). Discriminant and predictive validity of the WISC-III ACID profile among children with learning disabilities. *Psychology in the Schools*, *34*, 309-319. doi:10.1002/(SICI)1520-6807(199710)34:4<309::AID-PITS2>3.0.CO;2-G

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997b). Prevalence and diagnostic utility of the WISC-III SCAD profile among children with disabilities. *School Psychology Quarterly*, *12*, 235-248. doi:10.1037/h0088960

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children–Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003a). *Wechsler Intelligence Scale for Children–Fourth Edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003b). *WISC-IV technical and interpretive manual*. San Antonio, TX: Psychological Corporation.

Worrell, F. C., Vandiver, B. J., & Watkins, M. W. (2001). Construct validity of the Learning Behavior Scale with an independent sample of students. *Psychology in the Schools*, *38*, 207-215. doi:10.1002/pits.1011

Yen, C.-J., Konold, T. R., & McDermott, P. A. (2004). Does learning behavior augment cognitive ability as an indicator of academic achievement? *Journal of School Psychology*, *42*, 157-169. doi:10.1016/j.jsp.2003.12.001

Author Biographies

Gary L. Canivez is a professor of psychology at Eastern Illinois University, principally involved in the school psychologist training program. His research interests include psychometric reliability and validity investigations of measures of intelligence, achievement, psychopathology, and test bias.

Allison R. Gaboury is a school psychologist with Puyallup School District and earned her specialist in school psychology degree at Eastern Illinois University.