*Journal of Psychoeducational Assessment*
2005, *23*, 3-14

# ASSESSING THE CONSTRUCT VALIDITY OF THE ADJUSTMENT SCALES FOR CHILDREN AND ADOLESCENTS

Gary L. Canivez
*Eastern Illinois University*

Katie Sprouls
*Dysart Unified School District #89, El Mirage, Arizona*

This study examined the construct validity of the Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993). Distinct group differences and discriminative validity (Youngstrom, Findling, Danielson, & Calabrese, 2001) were examined with a sample of 106 students ranging from grade 1 through 6 (53 met the *DSM-IV/DISC-IV* criteria for ADHD, and 53 were members of a randomly selected and matched control group). Statistically significant group differences were observed with large effect sizes. Further, the ASCA demonstrated near perfect discrimination of the two groups and all diagnostic efficiency statistics were highly supportive in differentiating students meeting *DISC-IV/DSM-IV* criteria for ADHD from a random and matched control group

Current psychological assessment practice is illustrated by a growing preference for objective assessment techniques that can facilitate links between assessment and intervention (Piacentini, 1993; Reschly & Ysseldyke, 1995). Standardized behavior rating scales and checklists are the most frequently used instruments by school psychologists in assessing emotional and behavioral difficulties in youths (Stinnett, Havey, & Oehler-Stinnett, 1994) and have achieved popularity among applied psychologists (Hart & Lahey, 1999; Merrell, 1994a, 2003). Teachers have been considered to be among the most accurate adult raters of child behavior (Kamphaus & Frick, 1996) and appear to use a normative perspective in rating child behaviors (Piacentini, 1993) due to their observation of many students across time and contexts.

One teacher report behavior rating scale that is gaining empirical support is the Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993). The ASCA is a nationally normed behavior-rating instrument designed to assess psychopathology in youths aged 5 through 17 (kindergarten through grade 12). The ASCA defines psychopathology through multisituational expression of problem behaviors assessed by having raters

---

Data were obtained as part of Katie Sprouls' specialist in school psychology thesis completed at Eastern Illinois University.

A preliminary version of this study was presented at the 2003 Annual Convention of the National Association of School Psychologists.

Correspondence concerning this article should be addressed to Gary L. Canivez, Department of Psychology, 600 Lincoln Avenue, Charleston, IL 61920-3099. Dr. Canivez may also be contacted via e-mail at glcanivez@eiu.edu or the World Wide Web at <http://www.ux1.eiu.edu/~cfglc>. Information on the ASCA may be obtained at <http://www.ux1.eiu.edu/~cfglc/ASCA.html>.

indicate which specific behaviors typify the child in a variety of circumstances and contexts (McDermott, 1993, 1994). It consists of six core syndromes and two supplementary syndromes. The six core syndromes, which have been found to be reliable across gender, age, and race/ethnicity (McDermott, 1993, 1994), include Attention Deficit/Hyperactive (ADH), Solitary Aggressive-Provocative (SAP), Solitary Aggressive-Impulsive (SAI), Oppositional-Defiant (OPD), Diffident (DIF), and Avoidant (AVO). These six core syndromes also combine to form two composite (second-order) or overall adjustment indexes: Overactivity (ADH, SAP, SAI, and OPD syndromes) and Underactivity (DIF and AVO syndromes). Delinquency (DEL) and Lethargic-Hypoactive (LEH) make up the two supplementary syndromes and are reliable for certain subgroups in the population. Core syndromes, supplementary syndromes, and overall adjustment scales are reported as normalized *T* scores ($M = 50$, $SD = 10$) and percentiles.

Extensive evidence for ASCA score reliability and validity is presented in the ASCA manual (McDermott, 1994) and subsequent independent studies. Internal consistency estimates for the total standardization sample ranged from .68 to .86 for the six core syndromes and two supplementary syndromes and equaled .92 for the Overactivity scale and .82 for the Underactivity scale. Test-retest stability coefficients ($N = 40$) over a 30–school day interval ranged from .66 to .91 for the six core syndromes and from .75 to .79 for the Overactivity and Underactivity scales. No significant differences in mean *T* scores were observed across the retest interval. Canivez, Perry, and Weller (2001) also found significant stability for the ASCA overall adjustment scales, core syndromes, and supplemental syndromes over a 90-day retest interval. Stability coefficients ranged from .49 to .68 for the core syndromes, supplementary syndromes, and overall adjustment scales *T* scores, and mean changes were less than .8 raw score points. Canivez et al. (2001) also found significant stability for the ASCA syndromic profiles (kappa ranging from .24 to .59) *and* discriminant classifications (kappa = .35), two additional methods of score interpretation.

Interrater agreement for ASCA syndrome *T* scores has been reported (McDermott, 1994; Watkins & Canivez, 1997), with statistically significant correlations for the core syndromes and global adjustment scales and no statistically or clinically significant mean differences found between raters. Watkins and Canivez (1997) replicated the McDermott (1994) interrater agreement findings for the ASCA Overactivity, Underactivity, and core syndrome *T* scores (McDermott, 1993, 1994), with correlations ranging from .55 to .85. Canivez and Watkins (2002) reported statistically significant interrater agreement for ASCA Syndromic Profile classifications (kappa ranging from .39 to .68, $ps < .0001$), whereas Canivez, Watkins, and Schaefer (2002) reported statistically significant interrater agreement for ASCA Discriminant Classifications (kappa = .51, $p < .0001$).

Evidence of convergent and divergent (discriminant) validity of the ASCA has also been reported. McDermott (1993, 1994) found convergent validity coefficients ranging from .65 to .91 when comparing the ASCA and the Revised Conners Teacher Rating Scale (CTRS; Trites, Blouin, & Laprade, 1982). All

four of the ASCA overactive core syndromes were highly correlated with the CTRS Hyperactivity and Conduct Problem factors. The low to near zero correlations between the Overactive and Underactive core syndromes of the ASCA supported the divergent (discriminant) validity for these two dimensions (McDermott, 1993, 1994). Correlations between the ASCA and Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983) were statistically significant among similar psychological dimensions or constructs (McDermott, 1993, 1994). Canivez and Bordenkircher (2002) and Canivez and Rains (2002) found support for convergent and divergent (discriminant) validity in comparing the ASCA and the Preschool and Kindergarten Behavior Scales (PKBS; Merrell, 1994b) among randomly selected preschool, kindergarten, and first-grade children. Specifically, the ASCA Overactivity global adjustment syndrome and core syndrome scores were significantly and moderately to highly correlated with the similar PKBS Externalizing composite and subscale scores (*r*s ranging from .48 to .84). Divergent (discriminant) validity was observed with low to near zero correlations between the ASCA Underactivity global adjustment syndrome and related core syndrome scores and the PKBS Externalizing composite and subscale scores (*r*s ranging from -.19 to .25).

Additional divergent (discriminant) evidence of construct validity for the ASCA has also been reported. McDermott (1995) found low negative correlations (except one comparison) between the ASCA and the Differential Abilities Scale (DAS; Elliott, 1990). Correlations between the ASCA and DAS ranged from -.24 (ASCA ADH and DAS Spelling) to .10 (ASCA OPD and DAS Nonverbal Reasoning Ability), indicating that psychological adjustment as measured by the ASCA accounted for no more than 6% of the variability in ability or achievement as measured by the DAS. These findings were replicated in a study by Canivez, Neitzel, and Martin (in press) with comparisons to the WISC-III, K-BIT, and several individually administered achievement tests.

Exploratory factor analyses and confirmatory analyses reported by McDermott (1993, 1994) have indicated that the ASCA items are best explained by an eight-factor model with six factors (core syndromes) generalizing across gender, race/ethnicity, and age, and two factors (supplemental syndromes) appropriate for specific subgroups in the population. Factor analyses of the six core syndromes produced a two-factor solution (Overactivity and Underactivity), which appears similar to the two-dimensional model (conduct problem/externalizing vs. withdrawal/internalizing) of child psychopathology frequently obtained in the assessment literature (Achenbach, 1991; Achenbach & Edelbrock, 1983; Cicchetti & Toth, 1991; Merrell, 1994a, 1994b, 2003; Reynolds & Kamphaus, 1992, 2004; Quay, 1986). Core syndrome specificity estimates were also reported to be higher than error estimates and indicated that the separate core syndromes can be meaningfully interpreted (McDermott, 1994). Canivez (2004) also found that the core syndromes produced the same two-factor solution (Overactivity and Underactivity) in an independent sample of 1,020 children and adolescents. Oblique (Promax rotation) and orthogonal (Varimax rotation) solutions produced nearly identical structure coefficients, suggesting independent factors. Further, the Overactivity and Underactivity factors correlated only .08 in the oblique solution, highlighting the independence of these two dimensions.

McDermott (1994) and McDermott et al. (1995) showed that the ASCA core syndromes also demonstrated good discriminative validity (Youngstrom, Findling, Danielson, & Calabrese, 2001) and diagnostic accuracy (approximately 80% correct classification regardless of group or demographic characteristic) in differentiating students with emotional disturbance ($n$ = 150) from age-, gender-, race-, and grade level–matched normal students ($n$ = 150), as well as separate groups of learning-disabled ($n$ = 360), speech/language disabled ($n$ = 29), and gifted ($n$ = 60) students. Positive predictive power estimates (PPP = 80.6%) also exceeded a recommended standard (.75) for diagnostic tests (Landau, Milich, & Widiger, 1991).

The present study appears to be the only discriminative validity study of the ASCA. In order for the ASCA to be used for diagnostic purposes it is critical to explore further its discriminative validity and diagnostic utility.

This study examined two types of evidence for the construct validity of the ASCA. The first method was an investigation of distinct group differences (ADHD vs. Normal; Cohen & Swerdlik, 2002; Messick, 1995) as evidence for construct validity. Distinct groups that differ on a particular construct should differ on a test purporting to measure that construct. The second method was an investigation of discriminative evidence of construct validity using discriminant function analyses (Tabachnik & Fidell, 2001) with subsequent diagnostic efficiency statistics (Kessel & Zimmerman, 1993) to assess the diagnostic accuracy or diagnostic efficiency in discriminating students meeting *DISC-IV/DSM-IV* criteria for ADHD from random and matched students from the same classroom. In order for tests to be used for diagnostic purposes, they should be assessed for their classification accuracy. Diagnostic efficiency statistics including sensitivity, specificity, positive predictive power, negative predictive power, false positive rate, false negative rate, and overall correct classification should be routinely and systematically reported (Kessel & Zimmerman, 1993). Additionally, kappa, a statistic measuring the level of agreement beyond chance for nominal scale data (Cohen, 1960; McDermott, 1988), should be reported when comparing test results to a diagnostic standard. To facilitate accurate calculation and reporting, Canivez and Watkins (1996) presented a spreadsheet template (Canivez, 1994) patterned after the recommendations of Kessel and Zimmerman (1993).

## METHOD

### *Participants*

A total of 106 students from grades 1 through 6 in a large southwestern suburban school district participated in the present study. Students ranged in age from 6 to 11 years ($M$ = 9.09, $SD$ = 1.54). As is typical for students referred for screening for ADHD, male students outnumbered female students.

*ADHD group.* Of the 53 students in the ADHD group, 38 (71.7%) were male, 15 (28.3%) were female, 24 (45.3%) were Caucasian, 6 (11.3%) were Black/African American, and 23 (43.4%) were Hispanic/Latino. Grade distributions were as follows: grade 1 ($n$ = 7, 13.2%), grade 2 ($n$ = 10, 18.9%), grade 3 ($n$ = 8, 15.1%), grade 4 ($n$ = 8, 15.1%), grade 5 ($n$ = 6, 11.3%), and grade 6 ($n$ = 14, 26.4%). Special education characteristics of the ADHD group includ-

ed 36 (67.9%) students with no disability classification, 13 (24.5%) with specific learning disability, 3 (5.7%) with speech/language disability, and 1 (1.9%) with some "other" disability.

*RMC group.* Of the 53 students in the random and matched control (RMC) group, 38 (71.7%) were male, 15 (28.3%) were female, 24 (45.3%) were Caucasian, 6 (11.3%) were Black/African American, and 23 (43.4%) were Hispanic/Latino. Grade distributions were as follows: grade 1 ($n$ = 7, 13.2%), grade 2 ($n$ = 10, 18.9%), grade 3 ($n$ = 8, 15.1%), grade 4 ($n$ = 8, 15.1%), grade 5 ($n$ = 6, 11.3%), and grade 6 ($n$ = 14, 26.4%). Special education characteristics of the ADHD group included 38 (71.7%) students with no disability classification, 12 (22.6%) with specific learning disability, and 3 (5.7%) with speech/language disability.

### Instruments

*NIMH Diagnostic Interview Schedule for Children-Version IV.* The NIMH Diagnostic Interview Schedule for Children-Version IV (DISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) is a comprehensive structured interview that encompasses 36 mental health disorders for children and adolescents, using *Diagnostic and Statistical Manual of the American Psychiatric Association* criteria (*DSM-IV/DSM-IV-TR*; APA, 1994, 2000). The DISC-IV is a widely used and studied mental health interview that has been tested in both clinical and general populations (Johnson, Barrett, Dadds, Fox, & Shortt, 1999; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000).

*Adjustment Scales for Children and Adolescents.* The Adjustment Scales for Children and Adolescents (ASCA; McDermott, Marston, & Stott, 1993) is a standardized behavior rating scale that was normed on a representative national sample of 1,400 youths, blocked according to gender, age, and grade level. It is appropriate for use with children aged 5 through 17 (grades K-12). The ASCA contains 156 items, 97 that are scorable for dimensions of psychopathology and, based on factor analyses, are singularly assigned to one of six core syndromes (Attention Deficit-Hyperactive [ADH], Solitary Aggressive-Provocative [SAP], Solitary Aggressive-Impulsive [SAI], Oppositional Defiant [OPD], Diffident [DIF], and Avoidant [AVO]) or two supplementary syndromes (Delinquent [DEL] and Lethargic [LEH]). The core syndromes are combined to form two composite indexes: Overactivity (OVR; ADH, SAP, SAI, OPD syndromes) and Underactivity (UNR; DIF and AVO syndromes). Core syndromes, supplementary syndromes, and overall adjustment scales are reported as normalized $T$ scores ($M$ = 50, $SD$ = 10) and percentiles. In general, psychometric characteristics of the ASCA are acceptable and meet standards for both group and individual decision making (Canivez, 2001; Salvia & Ysseldyke, 1995).

### Procedure

Teachers in a southwestern suburban school district were informed of the opportunity to refer for screening students demonstrating behavioral problems or symptoms suggestive of ADHD through school-based prereferral inter-

vention teams. Once referred, the student's parent or primary caregiver was contacted about the prereferral intervention team and asked to complete the DISC-IV with the second author. A total of 53 students were identified as meeting the DISC-IV criteria for ADHD and thus comprised the "ADHD" group. A sample of 53 students from the same classrooms was selected at random and matched to the ADHD group on variables of age, gender, race, and disability and served as the random and matched control (RMC) group. Classroom teachers then completed the ASCA rating forms. The same teacher completed ASCA rating forms on both the referred student and the control group student. Teachers were blind to the DISC-IV results obtained from the parent interview.

### *Data Analyses*

To provide estimates of construct validity via distinct group differences, one-way MANOVA and subsequent univariate one-way ANOVAs were conducted to assess differences between the ADHD group and the random and matched control (RMC) group on the ASCA core syndromes. Effect sizes were estimated with Glass's Δ (Glass & Hopkins, 1996). To examine the diagnostic accuracy of the ASCA, direct discriminant function analyses (Tabachnick & Fidell, 2001) with subsequent diagnostic efficiency statistics (Canivez, 1994; Canivez & Watkins, 1996; Kessel & Zimmerman, 1993) were used.

## RESULTS

### *Distinct Group Differences*

A one-way MANOVA for differences between the ADHD group and the RMC group with the six ASCA core syndromes serving as dependent variables was statistically significant, Wilks's $\Lambda = .24$, $F = 52.45$, $p < .0001$. Subsequent one-way univariate ANOVAs were statistically significant for five of the six ASCA core syndromes. Table 1 presents results of the univariate ANOVAs, and Table 2 shows the descriptive statistics and effect size estimates for the ASCA core syndromes. The ADHD group obtained significantly higher scores than the RMC group on the ASCA ADH, SAP, SAI, OPD, and AVO syndromes. Effect sizes for the ADH, SAP, SAI, and OPD syndromes were large, whereas the effect size for the AVO syndrome was moderate (Cohen, 1992).

Table 1
*Univariate ANOVAs for Adjustment Scales for Children and Adolescents Core Syndromes*

| Core Syndrome | *SS* | *SS* Error | *MS* | *MS* Error | *F* | *p* | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| ADH | 8247.41 | 6380.72 | 8247.41 | 61.35 | 134.43 | .0001 | .56 |
| SAP | 6993.59 | 9713.47 | 6993.59 | 93.40 | 74.88 | .0001 | .42 |
| SAI | 7422.35 | 8151.81 | 7422.35 | 78.38 | 94.69 | .0001 | .48 |
| OPD | 2415.43 | 9588.53 | 2415.43 | 92.20 | 26.20 | .0001 | .20 |
| DIF | 69.77 | 9188.08 | 69.77 | 88.35 | 0.79 | .3760 | .01 |
| AVO | 900.76 | 8520.00 | 900.76 | 81.92 | 11.00 | .0010 | .10 |

Note.—MANOVA for Adjustment Scales for Children and Adolescents Core Syndromes: Wilks's $\Lambda = .24$, $F$ (6, 101) = 52.45, $p < .0001$, Multivariate Effect Size = .76, Power = 1.0. Univariate ANOVA $F$ tests $df$ (1, 104). On all significant effects, students with ADHD obtained higher ASCA scores than students in the control group. ADH = Attention-Deficit Hyperactive, SAP = Solitary Aggressive (Provocative), SAI = Solitary Aggressive (Impulsive), OPD = Oppositional Defiant, DIF = Diffident, AVO = Avoidant.

Table 2
*Descriptive Statistics, F, and Effect Size Estimates for Differences between the RMC and ADHD Groups*

| | RMC | | ADHD | | | |
|---|---|---|---|---|---|---|
| Core Syndrome | *M* | *SD* | *M* | *SD* | *F* | Δ |
| Attention Deficit-Hyperactive | 51.43 | 8.10 | 69.08 | 7.56 | 134.43 | 1.77 |
| Solitary Aggressive (Provocative) | 47.15 | 6.73 | 63.40 | 11.89 | 74.88 | 1.63 |
| Solitary Aggressive (Impulsive) | 50.19 | 8.11 | 66.92 | 9.54 | 94.69 | 1.67 |
| Oppositional Defiant | 52.21 | 8.93 | 61.75 | 10.23 | 26.20 | .95 |
| Diffident | 54.77 | 10.29 | 53.15 | 8.41 | 0.79 | .16 |
| Avoidant | 48.04 | 8.76 | 53.87 | 9.34 | 11.00 | .58 |

Note.—RMC = Random and Matched Control, ADHD = Attention Deficit Hyperactivity Disorder, Δ = Glass's Delta (Glass & Hopkins, 1996).

### *Discriminative Validity*

The direct discriminant function analysis[1] was statistically significant; Wilks's $\Lambda = .24$, $\chi^2(6) = 144.44$, $p < .0001$. Diagnostic efficiency statistics for the discriminant function analysis based on ASCA core syndromes are presented in Figure 1. The overall correct classification of 96% illustrated the high degree of diagnostic accuracy and near perfect separation of the ADHD and RMC groups. The significant kappa coefficient ($\kappa = .92$, $z = 9.53$, $p < .0001$) indicated an excellent level of agreement (Fleiss, 1981) and almost perfect agreement between the ASCA-based discriminant classification and *DISC-IV/DSM-IV* classifications of ADHD. Diagnostic efficiency statistics presented in Figure 1 also show very low proportions of false positive and false negative classifications and very high levels of sensitivity, specificity, positive predictive power, and negative predictive power.

## DISCUSSION

Results of the distinct group differences analyses were as expected, with large differences observed between the ADHD and RMC groups on the ASCA and mean scores for the RMC group very close to the scale *T* score means. Many of the differences observed between the ADHD and RMC groups had effect sizes that far exceeded minimum levels required for being classified as "large" (Cohen, 1992). It is also interesting to note that the mean score on the ASCA ADH syndrome for the ADHD Group was in the "maladjusted" range (McDermott, 1993, 1994) and mean scores for the SAP, SAI, and OPD syndromes were in the "at risk" range (McDermott, 1993, 1994). Impulsive aggression was also higher than provocative aggression, as might be expected among a group meeting ADHD criteria. All ASCA core syndrome scores for the RMC group and scores on the ASCA DIF and AVO syndromes for the ADHD group were in the "adjusted" range (McDermott, 1993, 1994). Differences between groups, however, are a necessary but not sufficient condition for diagnostic utility and use of a test. Diagnostic utility of a test requires investigation of indexes such as overall correct classification, sensitivity, specificity, positive predictive power, negative predictive power, false positive rate, and false negative rate

[1]Fisher's linear discriminant function equations based on the ASCA core syndromes are as follows: RMC = 1.236(ADH) + .498(SAP) + .451(SAP) + .254(OPD) + .743(DIF) + .205(AVO) - 87.427; ADHD = 1.556(ADH) + .695(SAP) + .638(SAP) + .291(OPD) + .769(DIF) + .232(AVO) - 133.511.

(Kessell & Zimmerman, 1993; Landau et al., 1991; Meehl & Rosen, 1955; Milich, Widiger, & Landau, 1987). Further, Landau et al. (1991) and Milich et al. (1987) have recommended the use of positive and negative predictive power as more meaningful indexes of diagnostic utility of a test.

| | | Diagnosis | | |
|---|---|---|---|---|
| | | Present | Absent | Total |
| Test | Positive | 50 | 3 | 53 |
| | Negative | 1 | 52 | 53 |
| | Total | 51 | 55 | 106 |

***Results***

| | |
|---|---|
| Sensitivity (True Positive Rate) = | 0.98 |
| Specificity (True Negative Rate) = | 0.95 |
| Positive Predictive Power = | 0.94 |
| Negative Predictive Power = | 0.98 |
| False Positive Rate = | 0.05 |
| False Negative Rate = | 0.02 |
| Overall Correct Classification (Hit) Rate = | 0.96 |
| Observed Agreement Po = | 0.96 |
| Chance Agreement Pc = | 0.5 |
| Kappa = | 0.92 |
| Standard Error of Kappa = | 0.0971 |
| Significance Test for Kappa Ho: k = 0 Z = | 9.47 |
| $p < 0$ | two-tail test |
| $p < 0$ | one-tail test |

**FIGURE 1. Diagnostic efficiency table for the Adjustment Scales for Children and Adolescents. ©1984 by Gary L. Canivez, Ph.D. All rights reserved.**

The present study is the first to independently investigate the diagnostic utility of the ASCA in classifying ADHD. The results clearly show that the ASCA was quite accurate in correctly differentiating those meeting the *DISC-IV/DSM-IV* criteria for ADHD from those in the random and matched control group and resulted in very few false positive and false negative classifications. The observation that overall correct classifications, sensitivity, specificity, positive predictive power, and negative predictive power were higher and false positive and false negative rates lower than in McDermott et al. (1995) and Forbes (2001) is in part the result that in the present study the ADHD group was compared to a group of random, matched, presumably normal, nonreferred students who differ considerably from students with significant symptoms of ADHD. The McDermott et al. (1995) study examined preexisting groups previously diagnosed with disabilities by multidisciplinary evaluation teams, whereas Forbes (2001) studied students clinically referred for assessment to determine if they had ADHD. Sensitivity and specificity estimates in the present

study are also higher than those found by Doyle, Ostrander, Skare, Crosby, and August (1997) in examining the diagnostic efficiency of the BASC-Parent Rating Scale and CBCL. Unfortunately, Doyle et al. (1997) did not report positive and negative predictive power for comparison. As with the Forbes (2001) study, Doyle et al. (1997) compared those referred and diagnosed to those referred but not diagnosed, which could partially account for the lower sensitivity and specificity estimates. The diagnostic efficiency statistics in the present study are also considerably higher than those reported by Doyle, Biederman, Seidman, Weber, and Faraone (2000) in examining the diagnostic utility of a battery of neuropsycholol gical tests in identifying ADHD.

The present study has limitations that qualify conclusions and generalizations. First, the sample was comprised of students from one school district in a southwest suburban geographic area and as such was not representative of the United States population. Thus, generalizability is limited and replication of this study in other geographic areas is recommended. Further, although the sample included a racially/ethnically diverse sample, no Asian American or Native American/American Indian children were included and generalization to these groups is not recommended. Replication with better representation of these groups is also recommended. Another consideration is that although the children met the *DISC-IV/DSM-IV* criteria for ADHD, it should not be assumed that they were "diagnosed" with ADHD. Such a diagnosis would require additional assessment information that went beyond the scope of this study and the prereferral resources of the authors. The use of the *DISC-IV/DSM-IV* criteria for ADHD was simply to have an independent measure of ADHD to serve as a criterion.

The present results clearly illustrate strong evidence of construct validity (distinct group differences and discriminative) for the ASCA. Of greatest importance was the diagnostic utility of the ASCA in correctly classifying students while maintaining impressively low rates of false positive and false negative classifications. As such, clinicians can be more confident in the use of the ASCA in their assessments of students with ADHD. Future studies should continue to examine the discriminative validity of the ASCA and examine the ability of the ASCA to differentiate not only ADHD from random and presumably normal students but also ADHD from other externalizing problems such as oppositional-defiant and conduct disorders (Meehl & Rosen, 1955). These discriminations would be a more stringent test of the discriminative validity of the ASCA. Should the ASCA be capable of differentiating ADHD from other externalizing disorders such as oppositional-defiant disorder and conduct disorder at levels comparable to this study, then their application in an actuarial classification would be more strongly advocated (Meehl, 1956; Meehl & Rosen, 1955) and it might not be necessary to include or recommend other more costly methods of psychological assessment that have not shown strong empirical support of discriminative validity (Doyle et al., 2000).

## REFERENCES


Achenbach, T. M. (1991). *Manual for the Teacher's Report Form and 1991 Profile.* Burlington, VT: University of Vermont Department of Psychiatry.

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

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

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: Text revision (DSM-IV-TR)* (4th ed.). Washington, DC: Author.

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. Impara & B. Plake (Eds.), *The fourteenth mental measurements yearbook* (pp. 22-24). Lincoln, NE: 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.

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

Canivez, G. L., Neitzel, R., & Martin, B. E. (in press). 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.*

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

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.

Canivez, G. L., & Watkins, M. W. (1996). Automated calculation of diagnostic efficiency statistics. *Behavior Research Methods, Instruments, & Computers, 28,* 132-133.

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.

Canivez, G. L., Watkins, M. W., & Schaefer, B. A. (2002). Interrater agreement of Discriminant Classifications for the Adjustment Scales for Children and Adolescents. *Psychology in the Schools, 39,* 375-384.

Cicchetti, D., & Toth, S. L. (1991). A developmental perspective on internalizing and externalizing disorders. In D. Cicchetti & S. L. Toth (Eds.), *Internalizing and externalizing expressions of dysfunction: Rochester symposium on developmental psychopathology* (Vol. 2, pp. 1-19). Hillsdale, NJ: Erlbaum.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Cohen, R. J., & Swerdlik, M. E. (2002). *Psychological testing and assessment: An introduction to tests and measurement* (5th ed.). Mountain View, CA: Mayfield.

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 & Clinical Psychology, 68,* 477-488.

Doyle, A., Ostrander, R., Skare, S., Crosby, R. D., & August, G. J. (1997). Convergent and criterion-related validity of the Behavior Assessment System for Children-Parent Rating Scale. *Journal of Clinical Child Psychology, 26,* 276-284.

Elliott, C. D. (1990). *Differential Ability Scales; Introductory and technical handbook.* San Antonio: The Psychological Corporation.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

Forbes, G. B. (2001). A comparison of the Conners' Parent and Teacher Rating Scales, the ADD-H Comprehensive Teacher's Rating Scale, and the Child Behavior Checklist in the clinical diagnosis of ADHD. *Journal of Attention Disorders, 5,* 25-40.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Needham Heights, MA: Allyn and Bacon.

Hart, E. L., & Lahey, B. B. (1999). General child behavior rating scales. In D. Shaffer, C. P. Lucas, & J. E. Richters (Eds.), *Diagnostic assessment in child and adolescent psychopathology* (pp. 65-87). New York: Guilford.

Johnson, S., Barrett, P., Dadds, M., Fox, T., & Shortt, A. (1999). The Diagnostic Interview Schedule for Children, Adolescents, and Parents: Initial reliability and validity data. *Behaviour Change, 3,* 155-164.

Kamphaus, R. W., & Frick, P. J. (1996). *Clinical assessment of child and adolescent personality and behavior.* Boston: Allyn and Bacon.

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.

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.

McDermott, P. A. (1988). Agreement among diagnosticians or observers: Its importance and determination. *Professional School Psychology, 3,* 225-240.

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

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

McDermott, P. A. (1995). Sex, race, class, and other demographics as explanations for children's ability and adjustment: A national appraisal. *Journal of School Psychology, 33,* 75-91.

McDermott, P. A., Marston, N. C., & Stott, D. H. (1993). *Adjustment Scales for Children and Adolescents.* Philadelphia: 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. *The Journal of Special Education, 29,* 337-354.

Meehl, P. E. (1956). Wanted–A good cookbook. *American Psychologist, 11,* 263-272.

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

Merrell, K. W. (1994a). *Assessment of behavioral, social, and emotional problems.* New York: Longman.

Merrell, K. W. (1994b). *Preschool and Kindergarten Behavior Scales.* Brandon, VT: Clinical Psychology Publishing.

Merrell, K. W. (2003). *Behavioral, social, and emotional assessment of children and adolescents.* Mahwah, NJ: Erlbaum.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741-749.

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.

Piacentini, J. (1993). Checklists and rating scales. In T. H. Ollendick & M. Hersen (Eds.), *Handbook of child and adolescent assessment* (pp. 82-97). Boston: Allyn and Bacon.

Quay, H. C. (1986). Classification. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (3rd ed., pp. 1-42). New York: Wiley.

Reschly, D. J., & Ysseldyke, J. E. (1995). School psychology paradigm shift. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-III* (pp. 17-32). Washington, DC: National Association of School Psychologists.

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

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

Salvia, J., & Ysseldyke, J. E. (1995). *Assessment.* Boston: Houghton Mifflin.

Shaffer, D., Fisher, P., Lucas, C., Dulcan, M., & Schwab-Stone, M. (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.

Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12,* 331-350.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.

Trites, R. L., Blouin, A. G. A., & Laprade, K. (1982). Analysis of the Conners Teacher Rating Scale based on a large normative sample. *Journal of Consulting and Clinical Psychology, 50,* 615-623.

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

Youngstrom, E. A., Findling, R. L., Danielson, C. K., & Calabrese, J. R. (2001). Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychological Assessment, 13,* 267-276.