Gender and Ethnic Differences in ADHD as Assessed by Behavior Ratings
Journal of Emotional and Behavioral Disorders 2000; 8; 38
DOI: 10.1177/106342660000800105

The online version of this article can be found at:
http://ebx.sagepub.com/cgi/content/abstract/8/1/38

Published by:
Hammill Institute on Disabilities

and

SAGE
http://www.sagepublications.com

Additional services and information for Journal of Emotional and Behavioral Disorders can be found at:

Email Alerts: http://ebx.sagepub.com/cgi/alerts

Subscriptions: http://ebx.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav
Attention-deficit/hyperactivity disorder (ADHD) is a common childhood disorder. Research suggests that ADHD is 4 to 9 times more frequent in males than females, and the possibility of underidentification in females and overidentification in males has been suggested as an explanation for these statistics. As part of the diagnostic process, teachers are frequently asked to complete behavior rating scales. There is a lack of empirical data concerning the extent to which gender differences are evident on such rating scales. This study investigated the use of the ADHD-IV Rating Scale—School Version, with male and female students from ages 5 to 18 years. Results suggest that the ADHD construct is consistent across gender; however, there are differences across gender and ethnicity. For Caucasian children, externalizing behaviors are most salient in terms of discriminating between males and females. Implications for research and practice are discussed.
The available research on gender differences has revealed some differences with regard to the referral process for boys and girls. Kashani and colleagues (1979) found that females with ADHD were usually referred for learning problems rather than behavior problems, whereas boys identified as hyperactive were more frequently referred for behavior problems than learning problems. This would be consistent with the research findings that girls with ADHD tend to have more comorbid internal manifestations of the disorder, whereas boys with ADHD have been noted to express more aggressive overt types of behavior (Breen & Altepeter, 1990; Brown, Abramowitz, Madan-Swain, Eckstrand, & Dulcan, 1989; deHaas, 1986; Gordon & Mettelman, 1994). Thus, gender-correlated behavioral patterns may be more frequently identified as ADHD in boys than in girls due to the frequency of disruptive classroom behavior exhibited by boys (Breen & Altepeter, 1990).

One important caution to keep in mind is that gender differences may vary as a function of the sample used. For example, Gaub and Carlson (1997) found a trend for greater severity of inattention among females, but comparable levels of hyperactivity in females and males, in a clinically referred sample. However, this pattern was not evident among nonreferred children.

It has been suggested that one cause for the gender disparity in referrals for ADHD may rest in the scales used as part of the diagnostic process, and the use of general norms as opposed to genderspecific norms (Barkley, 1996). Gender differences have been noted on rating scales used to assess children suspected of having ADHD. Based on information provided in the manual for the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992), for the standardization sample, males were rated one third to one half of a standard deviation higher than females on hyperactivity, attention problems, and other subscales. For this reason, the BASC manual includes separate norm tables for males and females, as well as a combined norm table, and recommends that the same-sex norms be used for clinical diagnosis in order to identify those children whose ratings are significant for both their age and gender. Similarly, DuPaul and colleagues (1997) reported significant differences across gender on composite inattention and hyperactivity scores based on normative data from the ADHD-IV Rating Scale-School Version (ARS; DuPaul, Power, Anastopoulos, & Reid, 1998). DuPaul et al. (1998) provided separate norms for males and females.

In contrast, Silverthorn and colleagues (1996) concluded that separate norms by gender were not warranted based on their finding that girls and boys with ADHD did not differ on measures of severity and that diagnostic cut-scores identified boys and girls with equivalent levels of impairment. They further argued that to use separate norms might artificially reduce the difference in prevalence rates for ADHD for girls and boys. Gordon (1996) argued that boys are more often identified with ADHD because they are more likely to demonstrate the severity of symptomatology to warrant this diagnosis; therefore, gender referencing is not appropriate. Alternatively, the possibility has been suggested that ADHD manifests differently in girls as compared to boys (Gordon, 1996; Gordon & Mettelman, 1994).

When considering the possibility of differences in ADHD symptomatology across gender, ethnicity is a factor that should also be considered. Serious concerns have been expressed regarding the assessment of ADHD with children from ethnic minorities (Bauermeister, Berrios, Jimenez, Acevedos, & Gordon, 1990) and the use of behavior rating scales with these children (Reid, 1995). However, ADHD among ethnic minorities remains an understudied area. Studies that have used behavior rating scales have reported significant differences between Caucasian (see Note) and African American students (e.g., DuPaul et al., 1997; Epstein, March, Connors, & Jackson, 1998; Jarvinen & Sprague, 1995; Lambert, Sandoval, & Sassone, 1978; Reid et al., 1998; Waechter, Anderson, Juarez, Landsdorpf, & Madrigal, 1979). However, it is uncertain whether these differences were due to real differences in behavior among groups, rater bias due to ethnicity and/or socioeconomic factors, or a combination of the two. Only one study to date has addressed the issue of rater bias. This study found evidence that ethnic status appeared to bias behavior ratings (Sonuga-Barke, Minocha, Tayor, & Sandberg, 1993). Because ethnic minorities constitute a large portion of the population, and because research has demonstrated convincingly that there are interethnic differences on behavior rating scales, we should attend to ethnicity as a factor in gender differences.

The purpose of this study was to explore the extent and nature of gender differences in ADHD across different ethnic groups (i.e., Caucasian and African American). Although clinical diagnosis requires multimethod, multisource information, we focused on teacher-completed rating scale information, because teacher ratings are often viewed as indicative of functioning in the school environment and are, thus, an important component in the assessment process (e.g., Barkley, 1989). Our exploration involved four interrelated analyses:

1. We analyzed descriptive statistics.
2. We performed a multivariate analysis to assess whether or not different patterns of item means existed across gender for two ethnic groups.
3. We used structural equation modeling to investigate whether the same two-factor construct of ADHD was appropriate for both males and females.
4. We conducted a discriminant function analysis to determine the extent to which items representing ADHD symptoms best separated children by gender.

Because of the possibility that gender differences may vary depending on symptom severity, we conducted analyses on two samples: one representative of unselected children and a second representing children with severe or numerous symptoms of ADHD.
Method

Instrumentation

Teachers completed the ARS. The ARS was selected because it reflects current DSM-IV diagnostic criteria for ADHD. This scale consists of 18 items—9 that address hyperactivity–impulsivity (HI) symptoms and 9 that address inattentiveness (IA) symptoms—directly adapted from the diagnostic criteria for ADHD specified in the DSM-IV. The ARS has demonstrated adequate reliability and criterion-related validity (DuPaul, Power, McGooey, Ikeda, & Anastopoulos, 1998).

In order to minimize possible bias due to response set, IA symptoms were designated as odd-numbered items, and HI symptoms were designated as even-numbered items. Teachers were instructed to select the single response for each item that best described the frequency of the specific behavior displayed by the target child over the past 6 months or since the beginning of the school year. The frequency for each item was delineated on a four-point Likert scale ranging from 0 (never or rarely) to 3 (very often), with the higher scores indicative of greater ADHD-related behavior. Teachers were asked to rate the behavior of two students from their class roster (e.g., third male and fifth female on the class roster). Ratings were completed between October and May in the 1994–1995 or 1995–1996 school years. Estimated return rates ranged from 50% to 95% (∑ = 85%) across school districts.

Participants

Participants for this study included 3,322 children and youth ages 5 to 18 (referred to as the Total group) taken from the normative sample of the ARS (DuPaul et al., 1997). The Caucasian (CA) sample consisted of 1,338 males and 1,298 females. The African American (AA) sample consisted of 376 males and 310 females. Because the total behavior rating scores decrease as age levels increase, the possibility of an age by gender dependency was tested for both CA and AA ethnic groups. The results suggested that ages were proportional across gender for both groups. AA group, χ²(12, N = 686) = 7.14, p = .84; CA group, χ²(12, N = 2636) = 12.39, p = .71.

A subsample of participants who might have been considered at risk for ADHD (referred to as the At-Risk group)—who scored above the 90th percentile on HI, IA, or both factors—was created. The 90th percentile was selected because it is the recommended cut-score (DuPaul, Power, Anastopoulos, & Reid, 1998). This resulted in an At-Risk group consisting of 218 CA males, 203 CA females, 64 AA males, and 53 AA females.

Participants were largely women (82.0%) and Caucasian (93.4%). Other ethnic groups represented among teachers included African American (5.4%), Hispanic (0.7%), Native American (0.1%), Asian American (0.1%), and other (0.4%). The majority of teachers (91.1%) were general educators; the remainder (8.9%) were special educators. No dependency was found between special education or general education status and gender for either ethnic group, AA group, χ²(1, N = 686) = .43, p = .51; CA group, χ²(1, N = 2636) = .04, p = .85.

Analysis

Descriptive Data. Descriptive data analysis followed the guidelines suggested by Bracken and Barona (1991). Descriptive statistics for each group were computed along with reliability coefficients (Cronbach’s alpha) for IA and HI factors. To assess the magnitude of difference across gender, effect sizes for each item were computed for both the CA and AA groups. Because effect sizes are a standardized measure, they allow for direct comparison of differences. Effect sizes were calculated by subtracting the female mean from the male mean, then dividing by the square root of the pooled variance.

Multivariate Analysis. To test for differences in item means across gender and ethnicity, a 2 (gender) × 2 (ethnicity) multivariate analysis of variance was performed. Separate analyses were conducted for the Total group and the At-Risk group.

Structural Equation Modeling. This analysis was conducted only for the Total group, because the number of AA males and females in the At-Risk group was too small. To test whether the two-factor (HI and IA) model of ADHD based on the DSM-IV was consistent across gender, structural equation modeling (SEM) was used to compare the factor structure, factor correlations, item loadings, and item uniqueness across gender, following the procedures (LISREL 8) suggested by Joreskog and Sorbom (1993) and Benson (1987).

The two-factor model was shown to adequately model observed data when factor analysis was conducted on the ARS during the norming process (DuPaul, Power, Anastopoulos, & Reid, 1998). To determine if the factor structure was invariant across genders, a four-step procedure was used. First, males and females were compared with all parameters (i.e., item factor loadings, factor correlations, and item uniqueness) constrained to be equal to those of the males. This step provides a baseline estimate of model fit. Second, separate estimates of item-factor loadings were computed for males and females. Third, separate estimates of item-factor loadings and factor correlations were computed for males and females. The last step involved computing the separate estimates of factor loadings, factor correlations, and item uniqueness (random measurement error and unique item variance) for males and females. The results from Steps 2 through 4 were then compared to the results of Step 1 to assess model fit with each additional freed parameter. The chi-square difference test was used to determine whether freeing a restraint improved the overall fit of the model (Bentler & Bonett, 1980). All parameter estimates were performed using covariance matrices and generalized least-square estimation.

Discriminant Function Analysis. To determine the extent to which ARS items were predictive of gender, we conducted discriminant function analysis using a procedure suggested by Huberty (1984). A step-wise analysis was used to determine the most parsimonious group of
predictor variables and to analyze each variable in terms of its effect on between-gender discrimination. Separate analyses were conducted for the Total group and the At-Risk group.

RESULTS

Descriptive Data

Item means for the Total group and the At-Risk group are shown in Figures 1 and 2. The reliability of the scale for all groups was consistently high. Cronbach's alpha for both the HI and IA scales was greater than .90 for AA and CA participants for both the Total and At-Risk groups. Table 1 shows between-gender effect sizes by item, separately for CA and AA ethnic groups within the Total group and the At-Risk group. In general, between-gender differences were slightly greater for the CA group than for the AA group. However, for some items such as Item 4 (leaves seat), this was not the case.

Multivariate Analysis

For most items, between-gender differences tended to be higher for the At-Risk group than for the Total group. However, in the case of the AA At-Risk group, between-gender differences decreased markedly for four of the nine IA items: 1 (fails to attend/careless), 5 (doesn't listen), 7 (fails to finish work), and 9 (difficulty with organization). As Table 1 shows, the average between-gender effect sizes for the Total group were CA = .40, AA = .36, and for the At-Risk group, CA = .52, AA = .34.

SEM Analysis

Tables 2 and 3 show the results of the SEM analysis. The results at Step one, where all parameters are invariant, show that all fit measures for both the CA (Table 2) and AA (Table 3) groups are at levels suggesting an acceptable fit across gender. The goodness of fit index is near the .90 level for both ethnic groups, suggesting an adequate fit. Also, the comparative fit index and the non-normed fit index are well above .90 for both groups. Additionally, the root mean square error of approximation is at or below the .08 level suggested as indicative of acceptable fit. This implies that, from a measurement perspective, there are no major qualitative distinctions across gender for either the CA or the AA group. When separate estimates of item uniquenesses were made, there was a statistically significant decrease in chi-square for both groups. However, because the fit at Step one, where all parameters were constrained across gender, is acceptable, the most conservative interpretation is that the same two-factor ADHD model holds across gender for both ethnic groups.
TABLE 2
Test of Model Fit Across CA Males and Females

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$T(df)$</th>
<th>$\chi^2(df)$</th>
<th>GFI</th>
<th>RMSEA</th>
<th>CFI</th>
<th>NNFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invariant</td>
<td>305</td>
<td>3111</td>
<td>—</td>
<td>—</td>
<td>.89</td>
<td>.083</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>+ Loadings</td>
<td>287</td>
<td>2986</td>
<td>18</td>
<td>25</td>
<td>.89</td>
<td>.084</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>+ Factor correlations</td>
<td>286</td>
<td>2985</td>
<td>1</td>
<td>1</td>
<td>.89</td>
<td>.084</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>+ Uniqueness</td>
<td>268</td>
<td>2368</td>
<td>18</td>
<td>617*</td>
<td>.91</td>
<td>.076</td>
<td>.99</td>
<td>.99</td>
</tr>
</tbody>
</table>

Note. GFI = goodness of fit index; RMSEA = root mean square error of approximation; CFI = comparative fit index; NNFI = non-normed fit index. *p < .01.

TABLE 3
Test of Model Fit Across AA Males and Females

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$T(df)$</th>
<th>$\chi^2(df)$</th>
<th>GFI</th>
<th>RMSEA</th>
<th>CFI</th>
<th>NNFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invariant</td>
<td>305</td>
<td>942</td>
<td>—</td>
<td>—</td>
<td>.84</td>
<td>.078</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>+ Loadings</td>
<td>287</td>
<td>916</td>
<td>18</td>
<td>26</td>
<td>.85</td>
<td>.080</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>+ Factor correlations</td>
<td>286</td>
<td>915</td>
<td>1</td>
<td>1</td>
<td>.85</td>
<td>.080</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>+ Uniqueness</td>
<td>268</td>
<td>832</td>
<td>18</td>
<td>83*</td>
<td>.86</td>
<td>.078</td>
<td>.99</td>
<td>.99</td>
</tr>
</tbody>
</table>

Note. GFI = goodness of fit index; RMSEA = root mean square error of approximation; CFI = comparative fit index; NNFI = non-normed fit index. *p < .01.

Discriminant Function Analysis

**CA Group.** For the Total group, a significant discriminant function for gender with eight predictors was obtained, $\chi^2(8, N = 2636) = 302.36, p < .0001$. Univariate statistics for each predictor are presented in Table 4. Group centroids for the male and female groups were .344 and -.354, respectively. For the At-Risk group, a significant discriminant function for gender with five predictors was also obtained, $\chi^2(5, N = 421) = 121.86, p < .0001$. Univariate statistics for each predictor are presented in Table 5. Group centroids for the male and female groups were .396 and -.471, respectively.

**AA Group.** For the Total group, a significant discriminant function for gender with four predictors was obtained, $\chi^2(4, N = 686) = 48.09, p < .0001$. Univariate statistics for each predictor are presented in Table 5. Group centroids for the male and female groups were .245 and -.297, respectively. For the At-Risk group, a significant discriminant function with two predictors was also obtained, $\chi^2(2, N = 117) = 19.67, p < .0001$. Univariate statistics for each predictor are presented in Table 5. Group centroids for the male and female groups were .396 and -.471, respectively.

**DISCUSSION**

The results of this study suggest that gender has a significant effect on teacher ratings of ADHD symptomatology. Results also suggest that there are no significant qualitative differences in symptomatology across gender and that the ADHD construct is consistent across gender. However, the results also suggest the possibility of cross-gender differences based on ethnicity.

Uniqueness of Symptomatology

If ADHD symptoms manifest differently across gender, then we would expect to see either a different pattern of item means and effect sizes, or differences in item variance/covariance structure. The results of this study suggest just the opposite. The observed pattern of item means was strikingly similar across gender and ethnicity. As Figures 1 and 2 show, there is a consistent pattern: the AA males were seen as most severe, the AA females and CA males were indistinguishable and distinctly separate from the AA males, and the CA females were seen as least severe and were distinctly separate from the AA females and CA males. The between-gender item effect sizes, which incorporate item variance for each group, further indicate that differences were consistent across items.

If females at risk for ADHD manifest higher rates of inattentive or hyperactive behavior, we would expect to see smaller between-gender item effect sizes for the At-Risk group than for the Total group for items on the IA or HI factors. However, this result was not generally obtained. Aside from smaller effect sizes for a few items for the AA At-Risk group, a significant discriminant function with two predictors was also obtained, $\chi^2(2, N = 117) = 19.67, p < .0001$. Univariate statistics for each predictor are presented in Table 5. Group centroids for the male and female groups were .396 and -.471, respectively.
(see Table 1), there appear to be no consistent pattern of item effect size differences between the Total group and the At-Risk group across either HI or IA factors.

Finally, the results of the SEM analyses suggest that, in terms of factor structure, factor loadings, and factor correlations, there is no difference between gender across ethnic groups. Thus, the results of both the descriptive analyses and the SEM analyses strongly suggest that, within the CA and AA groups, the ADHD construct was consistent across gender for the total sample.

We should caution that, because the number of AA students in the male and female At-Risk groups fell below the limits of a minimum of 100 participants per group (Loehlin, 1992), we did not use SEM to compare the At-Risk groups’ structures. Thus, the results of the SEM analyses are limited to the Total groups. It is possible that structural differences exist between males and females in the more severely involved groups.

## Effects of Gender on Teacher Ratings

The results of this study are consistent with previous analyses of the effects of gender on behavioral ratings (e.g., Reynolds & Kamphaus, 1992; Trites, Blouin, & Laprade, 1982) and strongly support the notion of gender differences in the perceived severity of symptom expression as assessed by teachers’ behavior ratings. The MANOVA results indicate significant cross-gender differences when one considers both the Total group and the At-Risk group. The observed between-gender effect sizes for items in Table 1 indicate that these differences are in the moderate range and exist for most scale items. This suggests that perceived gender differences are broad in nature, as opposed to being limited to a subset of items or a single dimension (i.e., IA or HI).

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Stepwise Discriminant Function Results for CA Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item #</td>
<td>Unique F</td>
</tr>
<tr>
<td>2</td>
<td>50.02</td>
</tr>
<tr>
<td>12</td>
<td>48.12</td>
</tr>
<tr>
<td>14</td>
<td>30.16</td>
</tr>
<tr>
<td>8</td>
<td>20.34</td>
</tr>
<tr>
<td>4</td>
<td>17.97</td>
</tr>
<tr>
<td>11</td>
<td>11.53</td>
</tr>
<tr>
<td>1</td>
<td>9.59</td>
</tr>
<tr>
<td>7</td>
<td>4.95</td>
</tr>
</tbody>
</table>

Note. Item 2 = fidgets; Item 12 = talks excessively; Item 14 = blurts answers; Item 8 = difficulty playing quietly; Item 4 = leaves seat; Item 1 = fails to attend/careless; Item 11 = avoids sustained mental effort; Item 7 = fails to finish work; Item 9 = difficulty with organization; Item 15 = easily distracted.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Stepwise Discriminant Function Results for AA Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item #</td>
<td>Unique F</td>
</tr>
<tr>
<td>13</td>
<td>4.17</td>
</tr>
<tr>
<td>17</td>
<td>7.69</td>
</tr>
<tr>
<td>2</td>
<td>6.07</td>
</tr>
<tr>
<td>7</td>
<td>10.72</td>
</tr>
</tbody>
</table>

Note. Item 13 = loses things; Item 17 = forgetful; Item 2 = fidgets; Item 7 = fails to finish work; Item 13 = loses things; Item 6 = runs/climbs excessively.
Furthermore, between-gender differences increased slightly in the At-Risk group; thus, gender differences increased as symptom severity increased. This is consistent with the fact that males are more likely to be identified as having ADHD than females. Interestingly, in the MANOVA analyses, when both gender and ethnicity are considered simultaneously, gender accounted for somewhat more variance than ethnicity, as evidenced by the $\eta^2$ statistics. Thus, the effect of gender appears to be equal to or slightly greater than that of ethnicity on teacher ratings.

In the discriminant function analyses, the items that most distinguished males from females in terms of a unique effect (i.e., when effects of correlation with other variables are parceled out) were not consistent across ethnic groups. For the CA group, a distinct pattern emerged. Items from the HI factor that reflect externalizing behaviors (i.e., fidgets, talks excessively, blurts out answers, difficulty playing quietly, and difficulty organizing tasks) constituted the majority of items and had the greatest effect on male-female separation for both the Total group and the At-Risk group. Thus, for the CA group, externalizing behavior seems to play a major role in defining gender differences. This is likely due, at least in part, to the fact that these behaviors are disruptive to the classroom environment and are more likely to be salient to the classroom teacher. For the AA group, no such pattern emerged. Fewer items attained significance, and only one item was significant for both the Total group and the At-Risk group (loses things necessary for tasks and activities).

This difference between the CA and AA groups appears to be the result of two factors. First, there is less difference between genders for the AA group, as evidenced by the between-gender effect sizes. Second, the shared variance for items appears to be greater for the AA group; thus, colinearity (i.e., high item intercorrelations) becomes a problem. This suggests that teachers’ perceptions of the ADHD-related behaviors of AA students is consistent across gender, with few distinguishing behaviors.

Why have some studies that have used behavior ratings (Breen, 1989; Breen & Altpeter, 1990; Horn et al., 1989; James & Taylor, 1990; Silverthorn et al., 1996) failed to find statistically significant between-gender differences? This is likely due to two reasons. First, sample sizes in previous studies were small, ranging from 39 (Breen, 1989) to 80 (Silverthorn et al., 1996). Thus, some studies may have lacked the statistical power needed to find gender differences of moderate size, such as those identified in the present study. A related problem lies in the small numbers of females with ADHD in the studies noted above, ranging from 13 (Breen, 1989) to 18 (James & Taylor, 1990). Consequently, the stability of results may be an issue. A second reason lies in the participant selection process. The present study used a randomized selection process from a school-based pop-
FIGURE 2. Item means for the At-Risk group.

ulation. In contrast, in studies that failed to find differences, the participants were drawn from clinically referred groups. One study used students identified by ICD-9 standards (James & Taylor, 1990), and another used students with pervasive ADHD (Horn et al., 1989). The reliance on clinically referred groups has been questioned on the grounds that children who were clinically referred may not be representative, as they may constitute an extreme group (Epstein, Shaywitz, Shaywitz, & Woolston, 1991).

It is also possible that there is a higher behavioral threshold for females (Eme, 1992). If so, females diagnosed as having ADHD would need to demonstrate extreme levels of behavior to receive an ADHD diagnosis. Therefore, clinically referred females with ADHD in these studies may not be representative of the general population of females with ADHD. However, we cannot rule out the possibility that there may be a subgroup of females who do not differ significantly from males.

Same or Different Norms

Present results have implications for the issue of separate ADHD norms for males and females and/or different ethnic groups. Generally stated, should we make judgments of deviance or disorder on the basis of an absolute standard or a relative standard? An absolute standard implies that a single threshold should be used to define behavior as disordered and that this threshold is valid across potentially mitigating factors, such as gender or ethnicity. A relative standard implies that there should be different thresholds for different groups (e.g., males and females or AA and CA groups) and that deviance or disorder should be judged within the context of group membership.

Results of the present study argue for the relative approach (separate standards for different groups) for using behavior rating scales. The use of an absolute standard implies the use of a single cut-score for all groups. This, in turn, would require pooling the scores of groups that are demonstrably different in terms of means and variances for all scale items. The result of this process is a "hybrid," composed of two dissimilar groups, with a cut-score that would not accurately represent either group. Males would be more likely to screen positive (exceed the cut-score), due to the inclusion of females; females would be less likely to screen positive, due to the inclusion of males. The fact that gender differences (i.e., between-gender effect sizes) were more pronounced as severity increased would further exacerbate this situation.

The problem of accurate representation is compounded when ethnicity is included. Our results demonstrated significant differences in item means across the AA and CA groups (males and females combined). Similarly, mean differences across AA and CA groups have been demonstrated in a number of previous studies and across different behavior rating scales (e.g., DuPaul et al., 1997; Epstein et al., 1998; Jarvinen & Sprague, 1995; Lambert et al., 1978; Reid et al.,
ences across gender and ethnic groups and the possibility of systematic rater in-
nicly different (Sonuga-Barke et al., 1998) and rater effects for students who are eth-
fects for AA students (Reid et al., 1998) also suggested the possibility of halo ef-
Rating Scale reflecting oppositional be-
ported that halo effects, which inflate be-
subject to halo effects and, thus, have
on the ARS rating scale. Parent ratings,
arguments are based on teacher ratings
the ARS rating scale. Parent ratings,
for example, may produce different re-
In addition, differences between
ethnic groups could be due to socioeco-
status. Future research should control for this variable. Finally, the
value of cut-scores should be determined by the relationships between cut-scores and functional impairment. Differences cannot be used synonymously with de-
viance. Further research will be required to determine the relationships among
gender, ethnicity, socioeconomic status, and functional impairment, as assessed
by behavior rating scales.

About the Authors

ROBERT REID is an associate professor in the department of Special Education and Communication Disorders at the University of Nebraska. His research focuses on children with attention and self-regulation difficulties. CYNTHIA A. RICCIO is currently an as-
sistant professor at Texas A & M University in College Station. Her primary research in-
terests include ADHD, neuropsychological assessment, and linguistic and cultural differences, as they impact assessment and differen-
tial diagnosis. ROBERT H. KESSLER re-
ceived his MA and EdS in school psychology from the University of Alabama. GEORGE
J. DUPOUL, PhD, is a professor and coordi-
nator of the school psychology program at
Lehigh University. Currently, he is investi-
gating the effects of early intervention and school-based interventions for students with
ADHD. THOMAS J. POWER, PhD, is an assistant professor of school psychology in pediatrics at the University of Pennsylvania and is co-director of the Center for Attention and Learning Problems at the Children's Hospital of Philadelphia. His research inter-
est includes assessment and intervention for children with ADHD and health promotion in urban schools. ARTHUR D. ANASTO-
POULOS, PhD, is an associate professor in the Department of Psychology at the University of North Carolina at Greensboro, where he also directs a specialty clinic serving chil-
dren, adolescents, and adults with ADHD. His research interests include the impact of
ADHD on family functioning. DIANA
ROGERS-ADKINSON is an associate pro-
essor and coordinator of the Emotional/
Behavioral Disorders Program at the Univer-
sity of Wisconsin-Whitewater. Her research
interests include the interaction of culture and
behavior and language function in children
with EBD. MARY-BETH NOLL has taught
students with EBD for 16 years and has
Taught in university teacher preparation pro-
grams for the past 9 years. Currently, she
is an associate professor and chairperson of the
Department of Special Education at St. Cloud
State University, where she continues to be
involved in planning a variety of state, re-
gional, and national professional develop-
ment opportunities for teachers of students
with EBD. Address: Robert Reid, 202 L
Barkley Center, University of Nebraska–
Lincoln, Lincoln, NE 68583-0732.

Note

For purposes of this study, Caucasian is defined
as persons of European-American ancestry.

References

Abikoff, H., Courtney, M., Pelham, W., &
Koplewicz, H. (1993). Teachers’ ratings of
disruptive behaviors: The influence of halo
effects. Journal of Abnormal Child Psy-
ology, 21, 519–533.
Diagnostic and statistical manual of men-
tal disorders (4th ed.). Washington, DC:
Author.
Arnold, L. E. (1996). Sex differences in ADHD:
Conference summary. Journal of Abnor-
mal Child Psychology, 24, 555–569.
Barkley, R. A. (1987). The assessment of at-
tention deficit-hyperactivity disorder. Be-
behavioral Assessment, 9, 207–233.


(continued from p. 37)


