Journal of Clinical Child and Adolescent Psychology
2002, Vol. 31, No. 4, 453–464

Sex Differences in Young Children Who Meet Criteria for Attention Deficit Hyperactivity Disorder

Cynthia M. Hartung
Oklahoma State University

Erik G. Willcutt
University of Colorado at Boulder

Benjamin B. Lahey
University of Chicago

William E. Pelham
State University of New York at Buffalo

Jan Loney
State University of New York at Stony Brook

Mark A. Stein
Children's National Medical Center, Washington, DC

Kate Keenan
University of Chicago

Examined sex differences in a mostly clinic-referred sample of 127 children (22 girls, 105 boys) who met Diagnostic and Statistical Manual of Mental Disorders (4th ed.; [DSM–IV], American Psychiatric Association, 1994) criteria for attention deficit hyperactivity disorder (ADHD) and 125 comparison children (24 girls, 101 boys) matched on age, sex, and race-ethnicity. Children in both groups ranged in age from 3 years, 10 months to 7 years, 0 months. Both girls and boys who met criteria for ADHD were more impaired than same-sex controls on a variety of measures when intelligence and other types of psychopathology were controlled. Teachers reported that boys with ADHD were more inattentive and more hyperactive/impulsive than girls with ADHD. These findings suggest that the diagnosis of ADHD is valid for both girls and boys in this young age range. Young girls and boys who meet DSM–IV criteria for ADHD are more similar than different, but boys tend to display more symptoms of ADHD, particularly in school.

The prevalence of attention deficit hyperactivity disorder (ADHD) has been found to be higher among boys and girls in many studies, although the male to female ratio is apparently smaller in population-based samples than in clinic samples (Barkley, 1998; Gomez, Harvey, Quick, Scharer, & Harris, 1999; Lahey, Miller, Gordon, & Riley, 1999; Szatmari, Offord, & Boyle, 1989). In addition, there is emerging evidence of sex differences among children who meet diagnostic criteria for ADHD. In a meta-analysis, Gaub and Carlson (1997) summarized the literature on sex differences in ADHD. To be included in the meta-analysis, each study was required to include (a) data comparing boys with ADHD to girls with ADHD, (b) participants age 13 and younger who had intelligence scores above 70, and (c) at least 10 participants in each group. As of 1994, Gaub and Carlson found 18 studies that met these criteria. Gaub and Carlson's meta-analysis showed no significant differences between boys and girls with ADHD on measures of impulsivity, academic performance, social functioning, or fine motor skills. In contrast, girls with ADHD scored lower than boys with ADHD on measures of global intelligence, whereas boys with ADHD exhibited higher levels of hyperactivity. Consistent with Gaub and Carlson's finding of higher levels of hyperactivity among boys who meet criteria for ADHD than girls, recent reviews by Carlson, Shin, and Booth (1999) and Milich, Balentine, and Lynam (2001) suggest that a greater of

Requests for reprints should be sent to Cynthia M. Hartung, 215 North Murray, Stillwater, OK 74078. E-mail: cmhart@okstate.edu

proportion of girls with ADHD meet criteria for the predominantly inattentive subtype than boys with ADHD.

Gaub and Carlson (1997) also found that boys with ADHD exhibit more comorbid psychopathology than girls with ADHD. These differences were dependent on the method of sample ascertainment, however. Nonreferred boys with ADHD exhibited higher levels of inattention, internalizing behaviors, peer aggression, and delinquent behaviors, whereas in clinic-referred samples boys and girls with ADHD were not significantly different on these variables. Two recent studies not in the Gaub and Carlson meta-analysis are also informative. Biederman et al. (1999) compared 140 clinic-referred girls with ADHD between the ages of 6 and 17 to 122 nonreferred girls matched on age and socioeconomic status. The girls who met criteria for ADHD exhibited higher levels of conduct, mood, and anxiety problems, received lower ratings of global functioning, and had more school problems than the comparison girls. In addition, a large population-based study by Reid et al. (2000) found that boys with ADHD are rated by teachers as exhibiting more conduct problems than girls with ADHD.

This growing literature on girls who meet criteria for ADHD is limited, however, by the fact that few studies included measures of functional impairment, and the studies that did assess impairment used only global measures of social and school functioning. Moreover, because studies that found girls with ADHD are more impaired than girls without ADHD did not control for intelligence scores and other types of psychopathology of children with ADHD, it is not clear that the greater levels of functional impairment are due to symptoms of ADHD or to other problems and deficits. One cannot conclude that the diagnosis of ADHD is valid for girls and boys until other possible reasons for functional impairment in girls with ADHD have been ruled out. For example, children with ADHD tend to be lower in intelligence than children without ADHD; it is possible, therefore, that children with ADHD are no more functionally impaired than other children with below-average intelligence. Alternatively, it is possible that ADHD symptoms are actually nonspecific indicators of other mental disorders and that ADHD symptoms do not create functional impairment beyond the impairment created by the other disorders. Such alternative hypotheses can be tested easily by studying the functional impairment associated with ADHD when intelligence and other types of psychopathology have been controlled, but only a few studies have done so to date (e.g., Lahey et al., 1998; Lavigne et al., 1996; Werry, Elkind, & Reeves, 1987). Therefore, in spite of many studies, more research examining the validity of the ADHD diagnosis for girls is needed.

This study extends the literature on sex differences in ADHD by examining a sample of young children who meet *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM–IV*]; American Psychiatric Association, 1994) criteria for ADHD. Specifically, the goal of this study was to test for sex differences in behavior and impairment in controlled analyses to examine the crucial question of potential sex differences in the validity of the diagnosis of ADHD. These analyses are based on the same sample used in an earlier report that did not consider sex differences (Lahey et al., 1998).

Method

Participants

One hundred twenty-seven children who met criteria outlined in the *DSM–IV* for one of the three subtypes of ADHD participated in the study. In addition, 125 children who did not meet criteria for ADHD and were mostly recruited from the same schools comprised a comparison group. Participants ranged in age from 3 years 10 months to 7 years 0 months with 98.4% being 4 to 6 years old on their last birthday. Twenty-two children with ADHD (17.3%) were girls and 105 (82.7%) were boys. Twenty-four comparison children (19.2%) were girls and 101 were boys (80.8%). All children with ADHD met the age of onset criterion for ADHD by virtue of their age, but to avoid circularity in analyses relating diagnoses to measures of functional impairment, the children were not required to meet the *DSM–IV* criterion of impairment in two or more settings.

Participants in the ADHD group were recruited in Chicago ($n = 58$) through a university child psychiatry clinic or Pittsburgh ($n = 69$) through either a university child psychiatry clinic (42%) or newspaper advertisements and flyers distributed to schools. Potential participants were consecutive referrals over a 2-year period and were referred to the study if (a) the presenting complaint was consistent with ADHD, (b) they lived with their biological mother, and (c) they did not exhibit pervasive developmental disorder, psychosis, or clear neurological disorder. Eight children who were recruited through newspaper advertisements were declared ineligible because the children had received diagnoses of pervasive developmental disorder, mental retardation, or seizure disorder. After the assessments were completed, an additional seven children were excluded from the sample because their intelligence scores were estimated to be less than 70.

Participants in the comparison group were recruited from the same schools, or from schools that served similar neighborhoods, as the participants in the ADHD group. According to mother reports, none had

ever been referred for mental health problems. Comparison children were selected from among those who volunteered to approximately match the children with ADHD in sex, ethnicity, and age. Two children who met criteria for ADHD were recruited as comparison children, and 12 children recruited for the ADHD group did not meet criteria for ADHD. These children were included in the ADHD and comparison group, respectively. This was viewed as the most conservative method of dealing with these participants, as it would tend to increase similarities between the ADHD and comparison group.

Procedures

The children and their biological mothers made one visit to the clinic for the assessment. Participants were told that the purpose of the study was to understand the difficulties experienced by young children who are considered to have ADHD in comparison to "average" children. Mothers gave written informed consent and the children gave oral assent to participate. Teachers completed rating scales via mail after the clinic visit. All participants were paid for their participation. The children and mothers were assessed and interviewed by two trained examiners with at least a bachelor's degree in education, psychology, or social work who were experienced in working with children. One examiner (Examiner 1) administered the cognitive testing to the child while a second examiner (Examiner 2) administered a structured self-report interview to the mother. Next, Examiner 1 administered the Diagnostic Interview Schedule for Children (version 2.3; DISC–2.3) to the mother while Examiner 2 administered the academic test and the self-report friendship questionnaire to the child. Immediately following the clinic visit, and prior to any contact with Examiner 2, Examiner 1 completed the Children's Global Assessment Scale (CGAS). Thus, Examiner 2 conducted the assessment of intelligence, achievement, and the child's self-report of social impairment with no knowledge of the child's symptoms of ADHD or other types of psychopathology. According to mother report, 20.2% of the boys and 9.1% of the girls in the ADHD group were being treated with psychotropic medication at the time of the assessment. However, mothers were instructed prior to the appointment not to administer psychotropic medication on the day of participation.

Measures

DISC–2.3. Biological mothers were asked to report symptoms of ADHD using the National Institute of Mental Health DISC–2.3 (Shaffer, Fisher, Piacentini, Schwab-Stone, & Wicks, 1993). Because the DISC–2.3 includes *Diagnostic and Statistical Manual for Mental Disorders* (3rd ed., revised [*DSM–III–R*]; American Psychiatric Association, 1984) criteria for ADHD, a supplementary module from the *DSM–IV* field trials (Lahey et al., 1994) that queried *DSM–IV* symptoms not in the DISC–2.3 was also administered. Previous studies (e.g., Schwab-Stone et al., 1996) have shown that test–retest agreement is good to excellent for ADHD as assessed by the DISC–2.3 (κ = .65 to .80). The interviewers in this study had similar levels of education and received the same training as in the Schwab-Stone et al. study. Because of the structured nature of the DISC interrater reliability has also been shown to be extremely high (κ = .95 for ADHD diagnosis; Hart et al., 1995; Lahey et al., 1995). Mothers were also asked to report on symptoms of oppositional defiant disorder (ODD) during the past 6 months, conduct disorder (CD) during the past 12 months, and anxiety disorders and depression during the past 6 months using the DISC–2.3. Schwab-Stone et al. reported that test–retest agreement is moderate to good for ODD and CD (κ = .56 to .73) and moderate for anxiety and depressive disorders (κ = .39 to .64). Teacher reports of numbers of ODD and CD symptoms since the beginning of the school year were obtained from the Disruptive Behavior Disorders (DBD) checklist.

DBD checklist. Teachers were asked to complete the *DSM–IV* version of the DBD (Pelham, Gnagy, Greenslade, & Milich, 1992). As in other studies (Milich, Hartung, Martin, & Haigler, 1993; Pelham et al., 1992), DBD items rated as occurring "pretty much" or "very much" were considered positive endorsements of a symptom. The internal reliability for ADHD symptoms in this sample was high for reports from the parent DISC (α = .96) and teacher DBD (α = .96). The diagnosis of ADHD was made by combining reports of symptoms from the parent DISC and teacher DBD using the "or rule," which considers each symptom of ADHD to be present if it is endorsed by either the parent or the teacher (Piacentini, Cohen, & Cohen, 1992). In addition, teachers completed the standardized IOWA Conners rating scale (Loney & Milich, 1982), which consists of items found to have both convergent and divergent validity in the measurement of hyperactive and oppositional-aggressive behavior.

Stanford–Binet Intelligence Scale (4th ed.) and Woodcock–Johnson Psychoeducational Battery (revised). The standard short-form of the Stanford–Binet Intelligence Scale (4th ed.; Thorndike, Hagan, & Sattler, 1986) was administered to obtain intelligence score estimates. In addition, subtests from the Broad Reading (i.e., Letter–Word Identification) and Math Reasoning (i.e., Applied Problems) composites of the Woodcock–Johnson Psychoeducational Battery (revised; Woodcock & Johnson, 1990) were administered to obtain estimates of academic achievement using age-based standard scores.

Peer preference scale. In addition to the DBD checklist, teachers were asked to complete several age-appropriate measures of social competence. First, teachers were asked to estimate the proportion of peers who "like," "dislike," and "ignore" the child (Dishion, 1990). We created a negative social preference score by subtracting the teacher's rating of the proportion of children who like the child from the proportion who dislike the child. That is, the widely used peer preference score (e.g., Melnick & Hinshaw, 2000) was calculated in reverse to invert the skew and allow Poisson regression analysis. To remove negative values for Poisson regression, the number 4 was added to all scores to make 0 the lowest score.

Teacher Assessment of Social Behavior (TASB). Teachers were also asked to complete the TASB (Cassidy & Asher, 1992), which provides ratings of four dimensions of social behavior (i.e., Prosocial, Shy/Withdrawn, Disruption, and Aggression). The internal reliability of the four scales is adequate to high (α = .62 for shy/withdrawn, .88 or higher for the other three scales; Cassidy & Asher, 1992). In a sample of kindergarten and first-grade children, children classified as low-accepted on the basis of peer sociometric ratings were found to have significantly lower scores on the prosocial behavior dimension and significantly higher teacher ratings of shy/withdrawn, disruptive, and aggressive social behavior (Cassidy & Asher, 1992).

Social Skills Rating System (SSRS). Finally, teachers were asked to complete the preschool version of the SSRS (Gresham & Elliott, 1990), which resulted in ratings of three social skills dimensions (i.e., Cooperation, Assertion, and Self-Control). The internal reliability of each SRSS scale is high (α = .93 to .94), and SSRS ratings are significantly associated with other measures of social competence and adaptive behavior (e.g., Flanagan, Alfonso, Primavera, Povall, & Higgins, 1996; Gresham & Elliott, 1990; Merrell, 1995).

Loneliness scale. Children completed a self-report instrument designed to assess ability to make and keep friends and feelings of being excluded from peer activities (Cassidy & Asher, 1992). Cassidy and Asher found that all but one of the 15 items from this measure loaded on a single factor and that the internal reliability of the measure was satisfactory (Cronbach's α = .79). Children with low ratings of social acceptance by peers were found to report significantly lower friendship scores than average or highly accepted children. Moreover, peers and teachers rated children with the lowest self-reported friendship scores as significantly less prosocial and more aggressive, and peers rated them as more shy (Cassidy & Asher, 1992). In this study, this measure was administered orally and the internal reliability in this sample was acceptable (α = .84).

Other measures of functional impairment. The CGAS (Setterberg, Bird, & Gould, 1992) is an instrument used to estimate global adaptive functioning that has been shown to have adequate test–retest reliability for the parent (r = .78) and interviewer (r = .75) versions (e.g., Bird et al., 1996). Scores range from 1 to 100, and for each decile the rater is provided with a phrase that describes functioning at that level. Raters are asked to provide the single number that best represents the child's lowest level of functioning during the past 6 months. The parent and Examiner 1 each completed the nonclinician CGAS. Parents also completed the parent version of the Children's Impairment Rating Scale (Fabiano et al., 1999) and teachers completed the teacher's version. The Impairment Rating Scale is an instrument designed to assess the degree to which a child's symptoms impact functioning and the child's need for treatment. This measure has been shown to have high internal consistency for the parent (α = .95) and teacher (α = .97) versions and adequate test–retest reliability for the parent (.50 to .70, $p < .001$) and teacher (.49 to .61, $p < .001$) versions (Fabiano et al., 2002). Finally, parents reported whether their child had ever sustained an unintentional injury of moderate to serious severity.

Data Analysis

The data were analyzed in a series of planned comparisons between pairs of the four groups (i.e., girls with ADHD, boys with ADHD, comparison girls, comparison boys). The distributions of some dependent variables were reasonably normal (i.e., age, intelligence, academic achievement, subscales of the SSRS, and the parent and interviewer versions of the CGAS), and comparisons were made using t tests or analysis of covariance. Because the distributions of other dependent variables were highly skewed and kurtotic, with the greatest proportion of youths receiving the lowest score (i.e., numbers of symptoms of ADHD, ODD, and CD; the teacher's negative social preference score; subscales of the TASB; the children's social dissatisfaction scores; and the parent's and teacher's overall rating of need for treatment), these comparisons were made using Poisson regression. When the mode of response distributions is at or near zero, the data cannot be successfully transformed to approximate normal distributions. On the other hand, because such highly skewed and kurtotic data approximate the Poisson distribution, regression models can be estimated in generalized linear models that are based on the Poisson distribution rather than the normal distribution (Kleinbaum, Kupper, Muller, & Nizam, 1998). Statistical tests in the Poisson regressions were based on the

χ^2 statistic using robust ("empirical") standard errors to adjust for any overdispersion from the Poisson distribution (Kleinbaum et al., 1998). To estimate group differences in CD in Poisson regression, one randomly selected comparison girl was assigned a value of 1 CD symptom to conservatively estimate differences, because no comparison girls exhibited any CD symptoms; this allowed the model to converge. Finally, the child's lifetime unintentional injuries was treated as dichotomous (0 = no injuries; 1 = one or more injuries) and comparisons were made in logistic regression.

In all comparisons of groups involving measures of functional impairment, other variables that distinguished children with ADHD from the comparison children and that might explain, in part, group differences in impairment were controlled (intelligence and numbers of internalizing, ODD, and CD symptoms) by entering them simultaneously in the regression analyses. Site (Pittsburgh vs. Chicago) was also tested as a covariate in all initial models to control for any differences in recruiting strategies, but was dropped from the final models because it had no significant impact on any result.

Balancing Type I and Type II Errors

The primary goal of this study is to examine sex differences in ADHD. Given the small sample of girls with ADHD ($n = 22$), the issue of balancing Type I and Type II errors is particularly germane. Controlling for experiment-wise error rates is a widely accepted procedure to decrease the probability of Type I errors (detecting associations that are not real). However, Type II errors (failing to detect associations) are also problematic in studies using small samples sizes. Jaccard and Guilamo-Ramos (2002) pointed out that

> it can be seen that statistical power tends to be low even for a single outcome variable for sample sizes typical of clinical child research and that applying an experimentwise error rate correction across multiple outcomes only exacerbates an already bleak situation in terms of Type II errors. In such situations one might decide not to invoke the experimentwise controls because the effect on statistical power is too severe and conceptually costly. (p. 133)

Furthermore, the issue of controlling for Type II errors is especially critical when attempting to study a disorder in the minority sex. It is often difficult to recruit adequate numbers of individuals of both sexes who meet criteria for a particular disorder when the disorder is overwhelmingly more common in one sex (Hartung & Widiger, 1998). In the case of ADHD, it is often difficult to recruit adequate numbers of girls who meet criteria for this disorder. Researchers who have attempted to study ADHD in girls are often faced with relatively low power for detecting sex differences. Therefore, a bias toward controlling for Type I errors at the expense of Type II errors can add to the difficulty of trying to understand how a disorder manifests in the minority sex.

As recommended by Jaccard and Guilamo-Ramos (2002), we grouped our variables into "families" and controlled for the experiment-wise error rate within, but not across, families. In addition, in an effort to balance the risk of Type I and Type II errors in the study, we have noted and cautiously interpreted findings that were initially significant but do not reach significance after the family-wise correction has been made.

Results

As shown in Table 1, girls and boys who met criteria for ADHD did not differ from either same-sex comparison children or from each other on age or family income. Boys with ADHD had significantly lower intelligence scores and academic achievement scores than did their same-sex comparisons (all $ps < .001$), but girls with ADHD did not. However, there were no significant differences between girls and boys who met criteria for ADHD on any of these variables. The significant findings for boys remained significant even after a family-wise error correction was made by grouping all of the demographic variables (i.e., for 21 comparisons $p = .05/21 = .002$).

Sex Differences in Symptoms of ADHD Among Children Who Met Criteria for ADHD

Girls and boys who met criteria for ADHD were not compared to same-sex comparison children on ADHD symptoms because they differed on these symptoms by definition. For inattention there was no significant difference between girls and boys with ADHD based on mother report (Table 2). In contrast, teachers reported that boys with ADHD displayed significantly more classroom symptoms of inattention ($p < .01$) than girls with ADHD. The magnitude of this difference was moderate, as assessed by Cohen's (1992) d ($d = .68$). For hyperactivity–impulsivity, there was no significant sex difference based on mother report. However, teachers gave boys with ADHD higher ratings than girls with ADHD on hyperactivity–impulsivity as measured by the DBD checklist ($p < .01$, $d = .76$) and the IOWA Conners rating scale ($p < .001$, $d = .85$). These significant sex differences based on teacher reports remained significant even after a family-wise error correction was made by grouping measures of ADHD (i.e., inattention and hyperactivity–impulsivity; for five comparisons, $p = .05/5 = .01$).

Table 1. *Descriptive Statistics and Planned Comparisons for Potential Covariates by Experimental Group and Sex*

	Boys				Girls				Boys	Girls	ADHD
	ADHD[a]		Comparison[b]		ADHD[c]		Comparison[d]		ADHD vs. Comparison	ADHD vs. Comparison	Girls vs. Boys
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*t*	*t*	*t*
Continuous variables											
Age	5.20	0.78	5.19	0.78	5.41	0.73	5.12	0.80	–0.11	–1.25	–1.16
Annual family income (1,000s)	39.63	34.14	48.06	31.02	41.27	37.57	49.11	31.02	1.93	1.33	–0.22
Intelligence score	91.42	12.92	101.9	14.41	94.54	14.00	98.88	13.71	5.50***	1.06	–1.02
Math age–standard score	98.34	17.17	111.88	16.72	99.50	14.67	106.62	18.09	5.72***	1.46	–0.29
Reading age–standard score	96.74	15.46	106.17	16.60	99.18	14.49	102.58	11.93	4.22***	0.87	–0.68
	%		%		%		%		χ^2	χ^2	χ^2
Categorical variables											
Race–ethnicity (%White)	63.8		64.4		68.2		62.5		0.01	0.16	0.15
Site (% Pittsburgh)	53.3		52.5		59.1		58.3		0.02	0.00	0.24

Note: ADHD = attention deficit hyperactivity disorder. Significance levels are presented without correcting for multiple comparisons.
[a]$n = 105$. [b]$n = 101$. [c]$n = 22$. [d]$n = 24$.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 2. *Descriptive Statistics and Planned Comparisons for Inattention and Hyperactivity–Impulsivity by Experimental Group and Sex*

	Boys				Girls				ADHD
	ADHD		Comparison		ADHD		Comparison		Girls vs. Boys
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Poisson χ^2
Inattention									
Mother DISC Report	5.69	2.84	0.46	1.00	6.09	2.54	0.33	0.64	0.38
Teacher DBD Checklist	4.92	3.05	0.31	0.93	2.86	3.03	0.17	0.56	7.80**
Hyperactivity–impulsivity									
Mother DISC Report	6.17	2.59	0.99	1.38	6.14	2.71	0.88	1.08	0.00
Teacher DBD Checklist	5.09	3.06	0.45	1.05	2.73	3.17	0.08	0.28	9.73**
Teacher IOWA Conner's	2.62	0.86	1.23	0.47	1.91	0.81	1.04	0.20	12.60***

Note: ADHD = attention deficit hyperactivity disorder; DISC = Diagnostic Interview Schedule for Children; DBD = Disruptive Behavior Disorder. Significance levels are presented without correcting for multiple comparisons.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3. *Descriptive Statistics and Planned Comparisons for Oppositional, Conduct, and Internalizing Symptoms by Experimental Group and Sex*

	Boys				Girls				Boys	Girls	ADHD
	ADHD		Comparison		ADHD		Comparison		ADHD vs. Comparison	ADHD vs. Comparison	Girls vs. Boys
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Poisson χ^2	Poisson χ^2	Poisson χ^2
ODD symptoms											
Mother DISC Report	2.99	2.54	0.57	1.21	3.32	2.42	0.21	0.51	56.75***	24.05***	0.31
Teacher DBD Checklist	2.28	2.43	0.31	1.08	0.95	1.70	0.08	0.28	33.51***	5.56*	5.49*
Teacher IOWA Conner's[a]	1.75	0.85	1.10	0.37	1.32	0.60	1.00	0.00	53.13***	8.58**	5.28*
CD symptoms											
Mother DISC Report	1.17	1.65	0.11	0.37	0.41	0.67	0.00	0.00	31.56***	4.51*	4.39*
Teacher DBD Checklist	0.78	1.45	0.08	0.40	0.36	0.95	0.00	0.00	15.68***	2.43	1.58
Internalizing symptoms											
Mother DISC Report[b]	6.89	4.98	3.03	2.81	6.91	5.95	3.00	2.64	46.58***	9.29**	0.00

Note: ADHD = attention deficit hyperactivity disorder; ODD = oppositional defiant disorder; DISC = Diagnostic Interview Schedule for Children; DBD = Disruptive Behavior Disorder; CD = conduct disorder. Significance levels are presented without correcting for multiple comparisons.
[a]IOWA Conner's oppositionality–aggression subscale. [b]Sum of DISC anxiety and depression.
*$p < .05$. **$p < .01$. ***$p < .001$.

Of the 105 boys who met criteria for *DSM–IV* ADHD according to combined parent and teacher reports of symptoms, 71 (67.6%) met criteria for Combined type, 25 (23.8%) met criteria for Predominantly Hyperactive–Impulsive type, and 9 (8.6%) met criteria for Predominantly Inattentive type. Of the 22 girls with ADHD, 13 (59.1%) met criteria for Combined type, 4 (18.2%) met criteria for Predominantly Hyperactive–Impulsive type, and 5 (22.7%) met criteria for Predominantly Inattentive type. Consistent with reviews of earlier studies that suggest that a larger minority of girls than boys meet criteria for the Predominantly Inattentive type (Carlson et al., 1999; Milich et al., 2001), logistic regression revealed a marginally significant result suggesting that girls were more often diagnosed with the Predominantly Inattentive type than boys in this sample (Wald $\chi^2 = 3.44$, $p = .06$, odds ratio = 3.1).

Sex Differences in Comorbid Disruptive and Internalizing Symptoms

As shown in Table 3, boys who met criteria for ADHD exhibited significantly more symptoms of ODD than comparison boys based on mother DISC, teacher DBD, and teacher IOWA Conners (ps < .001). In addition, girls with ADHD exhibited more symptoms of ODD than comparison girls based on mother DBD ($p < .001$), teacher DBD ($p < .05$), and teacher IOWA Conners ($p < .01$). According to mothers, there was no significant difference in ODD levels between boys and girls with ADHD. According to teachers, boys with ADHD exhibited significantly more ODD symptoms ($p < .05$, $d = .64$) and higher ratings on the Oppositional-Aggressive scale of the IOWA Conners ($p < .05$, $d = .59$) than girls with ADHD. After a fam-

ily-wise error correction was made by grouping all of the ODD variables (i.e., for nine comparisons, p = .05/9 = .006), only some of these findings remained significant. Specifically, the difference between girls with ADHD and same-sex comparisons on teacher-reported ODD from the DBD no longer reached significance. In addition, sex differences on both teacher-reported oppositionality variables in children with ADHD were no longer significant.

With regard to CD symptoms, boys who met criteria for ADHD exhibited significantly more symptoms than comparison boys based on mother and teacher (ps < .001) reports, whereas girls with ADHD exhibited more CD symptoms than comparison girls based on mother (p < .05), but not teacher, reports. In addition, boys with ADHD exhibited significantly more CD symptoms than girls with ADHD based on mother (p < .05, d = .66), but not teacher, reports. After a family-wise error correction was made by grouping the CD variables (i.e., for six comparisons, p = .05/6 = .01), only some of these findings remained significant. Consistent with the ODD findings, the differences involving girls with ADHD (n = 22), as compared with comparison girls and boys with ADHD, were no longer significant after controlling for Type I errors.

Mothers also reported that both girls (p < .01) and boys (p < .001) with ADHD exhibited significantly more internalizing symptoms (i.e., anxiety and depression) than same-sex comparison children. There was not a significant difference between girls and boys with ADHD on internalizing symptoms, however. These differences remained significant even after a family-wise error correction was made (i.e., for three comparisons, p = .05/3 = .02).

Validity of ADHD in Young Girls and Boys

To determine if the diagnosis of ADHD is associated with significant impairment in both sexes in this age range, girls with ADHD and boys with ADHD were compared to same-sex comparison children on measures of functional impairment, controlling for intelligence and other types of psychopathology reported by each informant. To make family-wise error corrections, these impairment variables were divided into three groups (i.e., peer relationship, behavioral control, and global functioning) as shown in Table 4. For peer relationship variables, boys who met criteria for ADHD exhibited significantly more impairment on all measures than did the comparison boys when intelligence and other psychopathology were controlled. These comparisons were all significant at p < .001 with the exceptions of self-reported Friendship Difficulties (p < .01) and TASB Withdrawn (p < .05). Girls who met criteria for ADHD exhibited more impairment on Negative Peer Preference (p < .01), SSRS Cooperation (p < .001), and TASB Prosocial (p < .05). There were no significant sex differences for children with ADHD on any peer relationship variables. After a family-wise error correction was made for the peer relationship variables (i.e., for 18 comparisons, p = .05/18 = .003), most of these findings remained significant. However, for ADHD versus comparison boys, TASB Withdrawn no longer reached significance, and for ADHD versus comparison girls, TASB Prosocial and Negative Peer Preference no longer reached significance.

For the behavioral control variables, boys with ADHD were significantly more impaired than comparison boys on SSRS Self-Control (p < .001), TASB Aggression (p < .01), and TASB Disruption (p < .001). The difference for unintentional injuries did not reach clinical significance (p = .07). Girls with ADHD were more impaired than comparison girls on only one of the four behavioral control variables (i.e., TASB Disruption, p < .01). In addition, TASB Disruption was the only behavioral control variable that showed a significant sex difference for children with ADHD (p < .05, d = .71). After a family-wise error correction was made for the behavioral control variables (i.e., for 12 comparisons, p = .05/12 = .004), most of these findings remained significant. However, the sex difference on TASB Disruption for children with ADHD no longer reached significance.

For the global impairment measures, boys with ADHD were significantly more impaired than comparison boys on all four measures (all ps < .001). Girls with ADHD were also more impaired than comparison girls at varying levels of significance (i.e., Mother CGAS and Teacher Treatment Need, p < .001; Mother Treatment Need, p < .01, Interviewer CGAS, p < .05). Teacher Treatment Need was the only global functioning variable that showed a significant sex difference for children with ADHD (p < .05, d = .74). After a family-wise error correction was made for the global functioning variables (i.e., for 12 comparisons, p = .05/12 = .004), most of these findings remained significant. However, for ADHD versus comparison girls, Interviewer CGAS no longer reached significance, and the sex difference on Teacher Treatment Need for children with ADHD no longer reached significance.

Discussion

In this study, teachers reported that young boys who met criteria for *DSM–IV* ADHD were more inattentive and hyperactive–impulsive than young girls who met criteria for ADHD, but there were no significant sex differences in ADHD symptoms according to maternal report. The finding that teachers rated boys with ADHD as more hyperactive and inattentive than girls with ADHD is consistent with the results of the meta-analysis of Gaub and Carlson (1997), but that

Table 4. *Descriptive Statistics and Planned Comparisons for Functional Impairment Measures by Experimental Group and Sex*

	Boys				Girls				Boys	Girls	ADHD
	ADHD		Comparison		ADHD		Comparison		ADHD vs. Comparison	ADHD vs. Comparison	Girls vs. Boys
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Poisson χ^2	Poisson χ^2	Poisson χ^2
Peer relationships											
Negative Peer Preference[a]	5.62	2.95	2.13	2.24	4.59	2.91	2.04	2.31	34.07***	7.44**	0.41
SSRS Cooperation[a]	10.16	3.77	15.56	2.61	11.68	2.71	16.29	2.94	60.58***	21.36***	1.61
SSRS Assertion[a]	8.96	3.58	12.84	3.71	10.09	4.03	13.00	3.76	26.33***	3.45	1.11
TASB Prosocial (reversed)[a]	5.62	2.95	2.13	2.24	4.59	2.91	2.04	2.31	33.75***	5.27*	0.35
TASB Withdrawn[a]	6.65	2.45	5.73	2.54	7.04	2.44	5.96	2.91	4.32*	0.96	0.36
Friendship difficulties[b]	11.03	8.08	5.99	5.47	8.77	7.28	6.08	5.29	9.74**	0.18	0.77
Disruptive behavior											
SSRS Self-Control[a]	9.48	4.04	15.09	3.62	11.86	4.39	14.42	3.84	34.01***	0.72	1.34
TASB Aggression[a]	7.96	3.78	4.55	2.69	6.36	4.05	4.21	2.60	9.88**	0.84	0.12
TASB Disruption[a]	14.01	4.15	7.72	3.56	10.86	4.68	6.50	2.75	56.24***	9.65**	5.05*
Unintentional injuries (%)[c]	31.4		10.9		13.6		8.3		3.31	0.74	2.47
Global functioning											
Interviewer CGAS Rating[a]	62.29	15.78	85.08	9.33	63.32	13.83	86.67	9.95	29.62***	6.75*	0.35
Mother CGAS Rating[c]	63.76	16.10	84.80	9.98	60.59	14.79	87.04	5.54	28.48***	22.15***	1.72
Teacher Treatment Need[a]	4.36	1.95	0.67	1.32	2.82	2.20	0.13	0.63	74.33***	10.79***	4.80*
Mother Treatment Need[c]	3.68	2.19	0.39	0.98	3.63	2.08	0.33	1.27	60.04***	10.24**	0.11

Note: ADHD = attention deficit hyperactivity disorder; SSRS = Social Skills Rating Scale–Teacher Report; TASB = Teacher Assessment of Social Behavior; CGAS = Children's Global Assessment Scale. Treatment Need is from the Children's Impairment Rating Scale. Means and proportions are presented without adjustment for covariates. Significance levels are presented without correcting for multiple comparisons.
[a]Controlling for intelligence and teacher's report of ODD and CD symptoms. [b]Controlling for intelligence, mother's report of ODD, CD, and internalizing symptoms, and teacher's report of ODD and CD symptoms. [c]Controlling for intelligence, mother's report of ODD, CD, and internalizing symptoms.
*$p < .05$. **$p < .01$. ***$p < .001$.

meta-analysis also found higher rates of parent-reported ADHD symptoms among boys with ADHD than girls with ADHD. One possible reason for the discrepancy between this study and Gaub and Carlson's is the younger age of our sample. The younger age of our sample might also explain why we did not find a significantly higher percentage of girls with ADHD to have the Inattentive subtype.

Both girls and boys who met criteria for ADHD were reported to exhibit significantly more internalizing symptoms than same-sex comparison children. In addition, boys with ADHD were reported to exhibit significantly more disruptive symptoms than same-sex comparison children. For girls with ADHD, the findings were generally similar, but significant findings on some variables did not survive family-wise error corrections. Consistent with Gaub and Carlson (1997) and Biederman et al. (1999), among children with ADHD, teachers reported more ODD symptoms in boys than girls, whereas mothers reported that boys exhibited more symptoms of CD than did girls. However, in this study these findings were no longer significant after family-wise error corrections were made. Nonetheless, the effect sizes were moderate, and the results suggest that additional research should be conducted to explore this issue further. There was no significant sex difference in internalizing symptoms in this study among children with ADHD. Consistent with our findings, Gaub and Carlson found that nonreferred boys with ADHD exhibited higher levels of internalizing behaviors than nonreferred girls with ADHD, but there was no sex difference in internalizing symptoms in clinic samples of girls and boys with ADHD. Again, the younger age of our sample might also explain why girls with ADHD did not show more internalizing symptoms than boys with ADHD.

There were no significant sex differences in cognitive and academic measures in this sample. This result is consistent with several previous studies that found no differences on cognitive measures between boys and girls with ADHD (Breen, 1989; Horn, Wagner, & Ialongo, 1989), but some studies suggest that girls with ADHD have lower intelligence scores and more learning problems than boys with ADHD (Brown, Madan-Swain, & Baldwin, 1991; Gaub & Carlson, 1997; James & Taylor, 1990). The examination of sex differences in the cognitive and academic profiles of girls and boys with ADHD is an area that has important implications for identification and treatment and will require further cross-sectional and longitudinal study to reconcile these discrepant findings from studies that used different age groups and ascertainment strategies.

The differential comorbidity hypothesis suggests that boys with ADHD may be referred more frequently than girls with ADHD because boys are more likely to exhibit comorbid disruptive behaviors. Although the higher prevalence of ADHD among boys in community samples suggests that this hypothesis cannot fully account for the male predominance (Lahey et al., 1999), it may help to explain why the male-to-female ratio is so much higher in clinic-based versus nonreferred samples. Within our sample, the male-to-female ratio among children with ADHD without a comorbid diagnosis of ODD or CD was 3:1, whereas the male-to-female ratio among children with a comorbid diagnosis of ODD or CD was 7:1. Thus, girls with ADHD were less likely than boys with ADHD to meet diagnostic criteria for a comorbid disruptive behavior disorder. This finding provides further support for the hypothesis that differential comorbidity may play a role in the higher prevalence of ADHD among boys in clinical samples by contributing to a referral bias.

Most important, the *DSM–IV* diagnostic criteria for ADHD were found to have concurrent validity for girls as well as boys in this age range, in the sense of identifying children with significantly higher levels of functional impairment when intelligence and other types of psychopathology were controlled. Furthermore, few sex differences were found in impairment associated with ADHD. The overall similarity of our results in girls and boys is consistent with other studies, which have implications for validity. For example, family studies have found similar rates of ADHD in the biological family members of boys and girls (e.g., Biederman, Faraone, Keenan, Knee, & Tsuang, 1990; Faraone et al., 2000) and Rhee, Waldman, Hay, and Levy (1999) reported similar heritability of ADHD symptoms in both sexes.

Considerably more research is needed on the topic of sex differences in ADHD, particularly studies of population-based samples that are free of referral biases and are large enough to provide sufficient statistical power to test for sex differences among the diagnostic subtypes of *DSM–IV* ADHD. Future studies should also compare the stability and longitudinal course of ADHD in girls and boys and evaluate the possibility that sex-based definitions of ADHD may more accurately identify girls with impairing problems of inattention and hyperactivity–impulsivity (Hartung & Widiger, 1998).

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.

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

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

Biederman, J., Faraone, S. V., Keenan, K., Knee, D., & Tsuang, M. (1990). Family-genetic and psychosocial risk factors in *DSM–III* attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 526–533.

Biederman, J., Faraone, S. V., Mick, E., Williamson, S., Wilens, T. E., Spencer, T. J., et al. (1999). Clinical correlates of ADHD in females: Findings from a large group of girls ascertained from pediatric and psychiatric referral sources. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 966–975.

Bird, H. R., Andrews, H., Schwab-Stone, M., Goodman, S., Dulcan, M., Richters, J., et al. (1996). Global measures of impairment for epidemiologic and clinical use with children and adolescents. *International Journal of Methods in Psychiatric Research, 6,* 295–307.

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

Brown, R. T., Madan-Swain, A., & Baldwin, K. (1991). Gender differences in a clinic-referred sample of attention-deficit-disordered children. *Child Psychiatry and Human Development, 22,* 111–128.

Carlson, C. L., Shin, M., & Booth, J. (1999). The case for *DSM–IV* subtypes in ADHD. *Mental Retardation and Developmental Disabilities Research Reviews, 5,* 199–206.

Cassidy, J., & Asher, S. (1992). Loneliness and peer relations in young children. *Child Development, 63,* 350–365.

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

Dishion, T. (1990). The peer context of troublesome child and adolescent behavior. In P. Leone (Ed.), *Understanding troubled and troubling youth* (pp.128–153). Newbury Park, CA: Sage.

Fabiano, G. A., Pelham, W. E., Gnagy, E. M., Kipp, H., Lahey, B. B., Burrows-MacLean, L., et al. (1999, November). *The reliability and validity of the Children's Impairment Rating Scale: A practical measure of impairment in children with ADHD.* Poster presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Toronto, Ontario.

Fabiano, G. A., Pelham, W. E., Gnagy, E. M., Waschbusch, D., Lahey, B. B., Chronis, A. M., et al. (2002). *A practical impairment measure for attention-deficit/hyperactivity disorder (ADHD): Psychometric properties of the impairment rating scale.* Manuscript in preparation.

Faraone, S. V., Biederman, J., Mick, E., Williamson, S., Wilens, T., Spencer, T., et al. (2000). Family study of girls with attention deficit hyperactivity disorder. *American Journal of Psychiatry, 157,* 1077–1083.

Flanagan, D. P., Alfonso, V. C., Primavera, L. H., Povall, L., & Higgins, D. (1996). Convergent validity of the BASC and SSRS: Implications for social skills assessment. *Psychology in the Schools, 33,* 13–23.

Gaub, B. A., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1036–1045.

Gomez, R., Harvey, J., Quick, C., Scharer, I., & Harris, G. (1999). *DSM–IV* AD/HD: Confirmatory factor models, prevalence, and gender and age differences based on parent and teacher ratings of Australian primary school children. *Journal of Child Psychology and Psychiatry, 40,* 265–274.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System: Preschool level.* Circle Pines, MN: American Guidance Service.

Hart, E. L., Lahey, B. B., Loeber, R., Applegate, B., Green, S. M., & Frick, P. J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology, 23,* 729–749.

Hartung, C. M., & Widiger, T. A. (1998). Gender differences in the diagnosis of mental disorders: Conclusions and controversies of the *DSM–IV. Psychological Bulletin, 132,* 260–278.

Horn, W. F., Wagner, A. E., & Ialongo, N. (1989). Sex differences in school-age children with pervasive attention deficit disorder. *Journal of Abnormal Child Psychology, 17,* 109–125.

Jaccard, J., & Guilamo-Ramos, V. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Issues and recommendations. *Journal of Clinical Child and Adolescent Psychology, 31,* 130–146.

James, A., & Taylor, E. (1990). Sex differences in the hyperkinetic syndrome of childhood. *Journal of Child Psychology and Psychiatry, 31,* 437–446.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E., & Nizam, A. (1998). *Applied regression analysis and other multivariate methods.* New York: Duxbury.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., et al. (1994). *DSM–IV* field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151,* 1673–1685.

Lahey, B. B., Loeber, R., Hart, E. L., Frick, P. J., Applegate, B., Zhang, Q., et al. (1995). Four-year longitudinal study of conduct disorder in boys: Patterns and predictors of persistence. *Journal of Abnormal Psychology, 104,* 83–93.

Lahey, B. B., Miller, T. L., Gordon, R. A., & Riley, A. (1999). Developmental epidemiology of the disruptive behavior disorders. In H. Quay & A. Hogan (Eds.), *Handbook of the disruptive behavior disorders* (pp. 23–48). New York: Plenum.

Lahey, B. B., Pelham, W. E., Stein, M., Loney, J., Trapani, C., Nugent, K., et al. (1998). Validity of *DSM–IV* attention-deficit/hyperactivity disorder for young children. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 695–702.

Lavigne, J. V., Gibbons, R. D., Christoffel, K. K., Arend, R., Rosenbaum, B., Binns, H., et al. (1996). Prevalence rates and correlates of psychiatric disorders among preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 204–214.

Loney, J., & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice (the IOWA Conners). In M. Wolraich & D. K. Routh (Eds.), *Advances in developmental and behavioral pediatrics* (Vol. 3, pp. 113–147). Greenwich, CT: JAI.

Melnick, S. M., & Hinshaw, S. P. (2000). Emotion regulation and parenting in AD/HD and comparison boys: Linkage with social behaviors and peer preference. *Journal of Abnormal Child Psychology, 28,* 73–86.

Merrell, K. W. (1995). Relationships among early childhood behavior rating scales: Convergent and discriminant construct validity of the Preschool and Kindergarten Behavior Scales. *Early Education and Development, 6,* 253–264.

Milich, R., Balentine, A. C., & Lynam, D. R. (2001). ADHD combined type and inattentive type are distinct and unrelated disorders. *Clinical Psychology: Science and Practice, 8,* 463–488.

Milich, R., Hartung, C. M., Martin C., & Haigler, E. D. (1993). Behavioral disinhibition and underlying processes in adolescents with disruptive behavior disorders. In D. Routh (Ed.), *Disruptive behavior disorders in childhood* (pp. 109–138). New York: Plenum.

Pelham, W. E., Gnagy, E., Greenslade, K., & Milich, R. (1992). Teacher ratings of *DSM–III–R* symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 210–218.

Piacentini, J., Cohen, P., & Cohen, J. (1992). Combining discrepant information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology, 20,* 51–62.

Reid, R., Riccio, C. A., Kessler, R. H., DuPaul, G. J., Power, T. J., Anastopoulos, A. D., et al. (2000). Gender and ethnic differences in ADHD as assessed by behavior ratings. *Journal of Emotional and Behavioral Disorders, 8,* 38–48.

Rhee, S. H., Waldman, I. D., Hay, D. A., & Levy, F. (1999). Sex differences in genetic and environmental influences on *DSM–III–R* attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 108,* 24–41.

Schwab-Stone, M. E., Shaffer, D., Dulcan, M. K., Jensen, P. S., Fisher, P., Bird, H. R., et al. (1996). Criterion validity of the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC 2.3). *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 878–888.

Setterberg, S., Bird, H., & Gould, M. (1992). *Parent and interviewer version of the Children's Global Assessment Scale.* New York: Columbia University.

Shaffer, D., Fisher, P., Piacentini, J., Schwab-Stone, M., & Wicks, J. (1993). *Diagnostic Interview Schedule for Children.* New York: Columbia University.

Szatmari, P., Offord, D. R., & Boyle, M. H. (1989). Ontario child health study. Prevalence of attention deficit disorder with hyperactivity. *Journal of Child Psychology and Psychiatry, 30,* 219–230.

Thorndike, R., Hagan, E., & Sattler, J. (1986). *The Stanford–Binet Intelligence Scale* (4th ed.). Chicago: Riverside.

Werry, J. S., Elkind, G. S., & Reeves, J. C. (1987). Attention deficit, conduct, oppositional, and anxiety disorders in children: III. Laboratory differences. *Journal of Abnormal Child Psychology, 15,* 409–428.

Woodcock, R. W., & Johnson, M. B. (1990). *Manual for the Woodcock–Johnson Psycho-educational Battery–Revised (WJ–R).* Chicago: Riverside.

Received December 20, 2001
Accepted April 23, 2002