

---

## EMPIRICAL ARTICLES

---

### **Are There Sex Differences in the Predictive Validity of *DSM-IV* ADHD Among Younger Children?**

**Benjamin B. Lahey**

*Department of Health Studies and Psychiatry, University of Chicago*

**Cynthia M. Hartung**

*Department of Psychology, Oklahoma State University*

**Jan Loney**

*Department of Pediatrics, State University of New York at Stony Brook*

**William E. Pelham**

*Department of Psychology, State University of New York at Buffalo*

**Andrea M. Chronis**

*Department of Psychology, University of Maryland*

**Steve S. Lee**

*Department of Psychology, University of California at Los Angeles*

*We assessed the predictive validity of attention-deficit/hyperactivity disorder (ADHD) in 20 girls and 98 boys who met the Diagnostic and Statistical Manual for Mental Disorders (4th ed., American Psychiatric Association, 1994) criteria for ADHD at 4 to 6 years of age compared to 24 female and 102 male comparison children. Over the next 8 years, both girls and boys who met criteria for ADHD in Year 1 exhibited more ADHD symptoms and impairment than same-sex comparison children. Effect sizes were consistently large, indicating that the diagnosis of ADHD at 4 to 6 years of age has predictive validity for both sexes. Both girls and boys with ADHD in Year 1 also exhibited higher levels of symptoms of conduct disorder, major depression, and anxiety disorders in early adolescence than same-sex comparison children, controlling levels of the same symptoms in Year 1. This indicates both substantial homotypic and heterotypic continuity for ADHD in both sexes, but significant interactions with time indicated that childhood ADHD predicts more steeply rising symptoms of anxiety and depression during early adolescence in girls than in boys.*

Since the publication of the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed. [*DSM-III*],

American Psychiatric Association, 1980), empirical research has increasingly shaped definitions of mental disorders. Little attention has been given, however, to the extent to which diagnostic definitions are valid for both sexes (Crick & Zahn-Waxler, 2003; Hartung & Widiger, 1998). Because of the

---

Correspondence should be sent to Benjamin B. Lahey, Department of Health Studies, University of Chicago, 5841 South Maryland Avenue, MC 2007, Room W262 Chicago, IL 60637. E-mail: blahey@uchicago.edu

fundamental importance of this issue, we address the predictive validity of the *DSM-IV* (4th ed., American Psychiatric Association, 1994) definition of attention-deficit/hyperactivity disorder (ADHD) for girls and boys. Emerging literature on sex differences in ADHD suggests that there are sex differences in ADHD, some of which could be related to differences in validity.

Previous studies of population-based samples document that boys are more likely to meet the criteria for ADHD than girls (Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). In addition, meta-analytic reviews by Gaub and Carlson (1997) and Gershon (2002) and subsequent studies (Hartung et al., 2002; Newcorn et al., 2001) found that girls who meet criteria for ADHD tend to score lower than boys on measures of cognitive ability and exhibit fewer total ADHD symptoms. There also is evidence that a greater proportion of girls than boys with ADHD meet criteria for the inattentive subtype (Carlson, Shin, & Booth, 1999; Milich, Balentine, & Lynam, 2001) and that girls with ADHD exhibit fewer comorbid conduct problems but more internalizing symptoms than boys with ADHD (Abikoff et al., 2002; Gaub & Carlson, 1997; Gershon, 2002; Levy, Hay, Bennett, & McStephen, 2005). These findings raise the possibility that girls and boys with ADHD exhibit somewhat distinct forms of the disorder. That is, girls with ADHD exhibit moderate levels of ADHD symptoms that are more strongly correlated with emotional problems than conduct problems, whereas boys exhibit higher levels of ADHD symptoms that are more strongly correlated with conduct problems. In further support of this hypothesis, several studies have found marked sex differences in the neurophysiological correlates of ADHD (Barry, Clarke, McCarthy, & Selikowitz, 2006; Baving, Laucht, & Schmidt, 1999; Hermens, Kohn, Clarke, Gordon, & Williams, 2005) and possible differences in the genetic and environmental factors associated with ADHD (Mick, Biederman, Santangelo, & Wypij, 2003; Rhee, Waldman, Hay, & Levy, 1999; Silberg et al., 1996). If sex differences in the fundamental nature of ADHD exist, then the predictive validity of the diagnosis, which reflects the developmental course of ADHD and associated impairment, may well differ between girls and boys.

The only previous study of the predictive validity of ADHD in girls and boys found that referred boys who met *DSM-IV* criteria for ADHD in childhood were less likely than referred girls with ADHD to have inpatient psychiatric hospitalization in adulthood (Dalsgaard, Mortensen, Frydenberg, & Thomsen, 2002). This was a small follow-back study, however, rather

than a prospective study. Within these limitations, their findings suggested possibly greater predictive validity of ADHD among girls than boys, perhaps related to differences in the adult sequelae of childhood ADHD. Much remains to be learned about this topic, however.

We based this study on a sample of children who met *DSM-IV* criteria for ADHD at 4 to 6 years of age and nonreferred comparison children matched for age, sex, and race-ethnicity (Lahey et al., 1998). At the time of the first assessment, the children who met symptom criteria for ADHD exhibited greater social and academic impairment and more unintentional injuries than comparison children, controlling for age, sex, race-ethnicity, socioeconomic status, intelligence, and concurrent symptoms of other forms of psychopathology (Lahey et al., 1998). These analyses take advantage of data from eight repeated assessments over 9 years to determine if *DSM-IV* diagnostic criteria exhibit predictive validity for both girls and boys, in the sense of the continuing symptoms and functional impairment over time. We tested for both homotypic continuity (i.e., childhood ADHD predicting future ADHD) and heterotypic continuity (i.e., childhood ADHD predicting future symptoms of other mental disorders). That is, we tested for multifinality in the outcomes of childhood ADHD (Cicchetti & Rogosch, 1996).

Our analyses controlled for several child and family characteristics that are confounded with the diagnosis of ADHD: family income, the child's intelligence, the number of Year 1 internalizing symptoms, and the number of Year 1 conduct problems. By including these controls, the longitudinal analyses tested the hypothesis that Year 1 ADHD per se predicts future functional impairment in both girls and boys, rather than other characteristics that are partially correlated with ADHD. That is, our analyses set an appropriately high standard for concluding that *DSM-IV* diagnostic criteria for ADHD have predictive validity when the diagnosis is made in young children. Within the limits of statistical models, the many covariates rule out the possibility that other characteristics of young children with ADHD could be responsible for any predictive validity. Previously, we reported that the diagnosis of ADHD exhibits predictive validity over 3 years in the same sample (Lahey et al., 2004), but sex differences in predictive validity were not tested.

When studying sex differences in disorders that are less prevalent in one sex, Type II errors (failure to reject the null hypothesis when there are population differences) are usually a greater concern than Type I errors (Crick & Zahn-Waxler, 2003). In this study, relatively few girls met criteria for ADHD.

Although the eight repeated measures increased statistical power, the relatively small number of girls in this sample provided only modest power to detect small sex differences. This means that failures to detect significant sex differences were not necessarily informative, as they could reflect inadequate power. Nonetheless, the detection of any moderate or large sex differences in the predictive validity of ADHD would be highly informative and point to the need for additional study.

## Method

### Participants

Two cohorts of 3.8- to 7.0-year-old children were recruited in consecutive years in Chicago and Pittsburgh. In Chicago ( $n = 116$ ), all probands were recruited from a university child psychiatry clinic where they presented with parent or teacher complaints of ADHD. In Pittsburgh ( $n = 128$ ), 42% of ADHD probands were recruited from a university child psychiatry clinic and 58% were recruited through advertisements. Probands recruited through advertisements did not differ significantly on demographic or impairment measures from those who presented at the psychiatry clinics in Year 1 (Lahey et al., 1998). Participants were required to be enrolled in structured educational programs: 36% preschool, 43% kindergarten, 21% first grade, and 1% second grade. All participants were required to live with their biological mothers. Five potential probands (two in Pittsburgh and three in Chicago) were excluded because they received clinical diagnoses of pervasive developmental disorder, mental retardation, or seizure disorder. Comparison children had never been referred for mental health problems but were not excluded if they met criteria for a mental disorder other than ADHD. They were recruited from the same schools as probands or from schools in similar neighborhoods and were selected to match probands in terms of sex, ethnicity, and age. Of the 310 eligible participants, 259 (Chicago  $n = 120$ ; Pittsburgh  $n = 139$ ) participated.

### Measures

**Diagnostic measures.** Independent interviews of the mother and child were conducted concurrently by two lay interviewers. The Stanford-Binet Intelligence Scale Short Form (Thorndike, Hagen, & Sattler, 1986) was administered in the first two assessments, and the scores were averaged to estimate intelligence. Four children were excluded from these analyses because their average intelligence scores were  $< 70$ . The Diagnostic Interview Schedule for Children (DISC; Shaffer, Fisher,

Piacentini, Schwab-Stone, & Wicks, 1993) was administered to the parent during each assessment to query *DSM-III-R* (3rd ed., rev., American Psychiatric Association, 1984) diagnostic criteria for ADHD, oppositional defiant disorder, conduct disorder (CD), simple phobia, social phobia, agoraphobia, separation anxiety, panic disorder, overanxious disorder, major depression, and dysthymia during the past 6 months. Questions used in the *DSM-IV* field trials version of the DISC (Lahey, Applegate, McBurnett, et al., 1994) to assess *DSM-IV* symptoms not in *DSM-III-R* also were asked. A parallel version of the DISC also was administered to youth (CD and depression modules in Years 6 through 9 and the anxiety disorders module in Years 7 through 9). In addition, teachers completed the *DSM-IV* version of the DBD Rating Scale (Pelham, Gnagy, Greenslade, & Milich, 1992) each year. Items rated *pretty much* or *very much* were scored as symptoms (Pelham et al., 1992).

Symptoms were based on multiple informants according to the age of the informant and the kind of symptoms. Parents and teachers have been shown to be reliable and valid informants on oppositional defiant disorder and ADHD (Hart, Lahey, Loeber, & Hanson, 1994). Youth over the age of 8 years have been shown to report reliably and validly on their anxiety, depression, and CD symptoms but not oppositional defiant disorder and ADHD (Bird, Gould, & Staghezza, 1992; Hart et al., 1994; Jensen et al., 1999). Therefore, as in the *DSM-IV* field trials (Lahey, Applegate, Barkley, et al., 1994; Lahey, Applegate, McBurnett, et al., 1994), ADHD and oppositional defiant disorder symptoms were considered present if reported by the parent or teacher; CD symptoms were considered present if reported by the parent, youth, or teacher; and depression and anxiety symptoms were considered present if reported by either the parent or youth using the "or" rule (Piacentini, Cohen, & Cohen, 1992).

In the initial assessment, impairment was assessed in two ways to address the *DSM-IV* requirement that impairment be present in at least two settings to make the diagnosis of ADHD. First, the parent was asked in the DISC if the child's ADHD symptoms had caused problems (a) at home or with friends or (b) at school. Second, parents and teachers completed the Impairment Rating Scale (Fabiano et al., 2006), in which the child's need for treatment in a variety of areas is rated using 7-point scales ranging from 0 (*no problem; definitely does not need treatment*) to 6 (*extreme problem; definitely needs treatment*). Parents rated the child's need for treatment related specifically to problems with peers, siblings, and

parents; academic progress at school; self-esteem; and impact on the family. Teachers rated the child's need for treatment related specifically to problems with classmates and teachers, academic progress, classroom disruption, and self-esteem. In addition, parents and teachers both rated the child's overall need for treatment. In two samples, Impairment Rating Scale ratings of  $\geq 3$  on at least one scale optimally differentiated clinic and non-clinic children for both parents and teachers (Fabiano et al., 2006). Test-retest stability (different teachers 1 year apart) for the six Impairment Rating Scale scales was  $r = .39$  to  $.63$  ( $p < .001$ ; Fabiano et al., in press).

Children were said to exhibit ADHD if they met full *DSM-IV* criteria for symptoms, age of onset, and cross-situational impairment for any subtype of ADHD based on the same algorithm employed in a previous report from this sample (Lahey, Pelham, Loney, Lee, & Willcutt, 2005). This algorithm is based on both general concerns about the cross-situational impairment criteria and specific concerns about its application to younger children. In general, although it is essential that impairment or distress be present to give any mental health diagnosis, it is not clear why children who meet symptom criteria for ADHD and are significantly impaired in one setting would not be considered to exhibit the disorder (and qualified for third-party support of treatment). ADHD is the only disorder for which impairment is required in more than one setting. In addition, because children in this study were young at the time of the initial assessment (79% were in preschool or kindergarten in Year 1), it is likely that the full extent of their impaired functioning in home and particularly in school may not have been evident yet in the initial assessment. Therefore, functional impairment associated with ADHD, as defined previously, was required in either home or school or both during the initial assessment, but cross-situational impairment was defined in a developmentally sensitive manner. To avoid excluding children whose impairment across both settings only emerged when the demands of elementary school increased, children who met other criteria for ADHD were allowed to meet the requirement of cross-situational impairment by exhibiting impairment in both home and school during any one of the eight annual assessments. Using this strategy, 81% of the children who were said to meet full criteria for ADHD in this study exhibited cross-situational impairment in Year 1 and 98% did so by Year 4.

Using this algorithm, 20 girls and 98 boys met diagnostic criteria for ADHD in Year 1. The symptoms and impairment over time were compared with 24 girls and 102 boys who did not meet

symptom criteria for ADHD in Year 1. These analyses do not include 11 children who met symptom criteria for ADHD in Year 1 but did not exhibit cross-situational impairment in any assessment. This exclusion did not bias the analyses because essentially the same percentage of girls (4.4%;  $n = 2$ ) and boys (4.3%;  $n = 9$ ) were excluded on this basis. When assessing the proportion of children who continued to meet criteria for ADHD over each successive wave, they were similarly said to meet full criteria if they met symptom criteria, exhibited impairment in at least one setting during that wave, and exhibited impairment in both home and school during at least one wave.

**Adaptive functioning outcome measures.** Child adaptive functioning during Years 2 through 9 was assessed using a number of impairment measures that were not used to make the diagnosis of ADHD. The nonclinician version of the Children's Global Assessment Scale (CGAS; Setterberg, Bird, & Gould, 1992) was used to obtain ratings of the child's lowest level of overall functioning during the past 6 months. CGAS ratings were collected independently from the mother and interviewer who administered the DISC to the mother. In each year, the parent also was asked if the child had suffered an injury more serious than a scratch, bruise, or bump on the head that the parent attributed to the child's carelessness, impulsiveness, or poor judgment.

The child's regular-education teacher estimated the proportion of classmates who liked and disliked the child using a 5-point scale (Dishion, 1990). For each descriptor, the scale ranged from 1 (very few: less than 25%) to 5 (almost all: more than 75%). A social preference score (Coie, Dodge, & Coppotelli, 1982; Sandstrom & Cillessen, 2003) was created by subtracting *dislike* from *like* ratings. This index was scored as a negative social preference score (indicating greater disliking ratings relative to liking ratings) by reversing scores to approximate Poisson distributions for statistical analysis. The sum of two items (each rated 0 to 2) from the Social Skills Rating Scale (Gresham & Elliott, 1990) quantified teacher perception of classroom academic functioning: "produces correct schoolwork" and "finishes class assignments within time limits." Findings based on this measure should be interpreted cautiously because the wording of some ADHD symptoms (i.e., "makes careless mistakes in schoolwork" and "fails to finish schoolwork") imply impaired classroom academic functioning. Nonetheless, if children who met criteria for ADHD were not impaired on this measure during at least part of their school careers, it would raise questions

about the validity of ADHD. This is because impaired academic classroom functioning is one of the strongest justifications for diagnosing and treating ADHD in school-age children.

All symptom and impairment measures used in these analyses have been shown to have at least adequate reliability in this sample (Lahey et al., 2004, 2005).

### Data Analysis

These analyses are based on data from annual diagnostic assessments conducted over 9 years, except that no assessment was completed in Year 5. The average ages of the children in each of the successive eight assessments were 5, 6, 7, 8, 10, 11, 12, and 13 years. For the main outcome variables, three sets of planned comparisons were conducted between (a) girls who did and did not meet criteria for ADHD in Year 1, (b) boys who did and did not meet criteria for ADHD in Year 1, and (c) girls and boys who met criteria for ADHD in Year 1. The response variables (ADHD symptoms and diagnosis, adaptive functioning, and symptoms of co-occurring mental disorders) were assessed during Years 2 through 9. The first two sets of planned comparisons addressed the predictive validity of ADHD in girls and boys compared to same-sex comparison children, and the third set of comparisons assessed potential sex differences between girls and boys who met criteria for ADHD during the initial assessment.

Numbers of symptoms of each type of psychopathology and teacher ratings of classroom work and social preference were analyzed using log linear regression, specifying Poisson distributions in generalized estimating equations (Zeger & Liang, 1986). All longitudinal analyses specified autoregressive correlation structures and used the  $z$  statistic and robust standard errors. Generalized estimating equations has become a standard method for the analysis of correlated longitudinal data, partly because it is appropriate for the analysis of response variables with the highly skewed distributions that are common in psychopathology research and partly because it is efficient for the analysis of correlated data lacking homogeneity of variance (Dahmen, Rochon, König, & Ziegler, 2004). In addition, generalized estimating equations is based on fewer modeling assumptions than other longitudinal approaches, such as mixed models (Zeger & Liang, 1986). This is important because violations of assumptions are more problematic in small samples, such as in the tests of sex differences in this study. These advantages come at the cost of some loss of statistical power relative to some other statistical approaches.

The continuous CGAS ratings approximated normal distributions and were analyzed using linear regression in generalized estimating equations. Differences in rates of meeting full criteria for ADHD during Years 2 through 9 and other dichotomous response variables that were measured in each assessment wave were assessed using longitudinal binomial regression. Response variables defined by at least one occurrence (unintentional injury and special education placement) were assessed using logistic regression. In all figures, means are presented for Year 1 as a standard of reference, but all analyses of the outcomes were based only on Years 2 through 9 to address predictive validity. Interactions between time (year of assessment) and group were always tested but were reported only when significant at  $p < .05$ .

**Statistical controls.** All longitudinal tests of group differences in symptoms and impairment over Years 2 through 9 controlled for the child's age in Year 1 and time (years). Controlling for the child's age in Year 1 minimized any variance due to initial differences in age, and controlling for time allowed assessment of developmental change and proper variance estimates for testing group differences (i.e., differences between children with and without ADHD and between the sexes). Two fixed-design characteristics also were controlled in all outcome analyses: Pittsburgh or Chicago site and Cohort 1 or 2. These controls removed any extraneous variance due to site and cohort differences from the error term to increase power to detect group differences. Preliminary analyses were conducted to determine if there were sex differences in changes in the sample over time that required additional statistical controls. These variables included attrition from the sample, lack of participation of relevant informants, participation in psychosocial treatment, and use of prescribed psychoactive medication.

**Potential attrition biases.** The percentage of youth participating in assessments after Year 1 ranged from 93% to 96% in Years 2 through 6 and 90% to 91% in Years 7 through 9. Although some youth did not participate in every assessment, few dropped out entirely. Only 6.2% of the 244 youth included in the analyses were not assessed in either Year 8, Year 9, or both. The numbers of girls who met full criteria for ADHD in Year 1 who were assessed in each year ranged from 15 to 20, with all 20 of the girls with ADHD assessed in at least one of the last two years (Year 8 or Year 9). Similarly,

the numbers of comparison girls assessed in each year ranged from 20 to 24, with 21 comparison girls assessed in Year 8, Year 9, or both. The numbers of boys with ADHD in each assessment was 88 to 98, with 92 assessed in Year 8, Year 9, or both. The numbers of comparison boys in each assessment ranged from 92 to 102, with 96 in the last two assessments.

Rates of participation during Year 8, Year 9, or both did not differ for girls (93.2%) and boys (94%),  $p = .74$  (Fisher's exact probability test), indicating little chance that differential attrition biased tests of sex differences. There also were no significant differences at the  $p < .05$  level between children who did or did not participate in Year 8, Year 9, or both in their age at the time of the initial assessment, race-ethnicity, intelligence, family income, CGAS ratings, or numbers of symptoms of inattention, hyperactivity-impulsivity, or conduct problems. However, Poisson regression revealed that the 15 girls and boys who did not participate in the last two waves had more internalizing symptoms than youth who did participate in these waves,  $\chi^2 = 5.82$ ,  $p < .02$ . Furthermore, Year 1 internalizing symptoms predicted higher levels of anxiety symptoms,  $\chi^2 = 5.79$ ,  $p < .0001$ , and depression symptoms,  $\chi^2 = 2.79$ ,  $p < .02$ , during Years 6, 7, 8, and 9. This suggests that these findings may underestimate levels of such symptoms during the last assessments. Fortunately, because there was no differential attrition for girls and boys, it does not compromise the relative comparisons of the two sexes.

**Informant participation and blind interviews.** The percentage of children whose parent was interviewed in each successive assessment was 100%, 96%, 94%, 94%, 91%, 87%, 90%, and 89%, respectively. The percentage with completed youth assessments in each year was 100%, 94%, 91%, 91%, 89%, 85%, 87%, and 89%, respectively. The percentage with completed teacher assessments in each year was 98%, 88%, 83%, 84%, 81%, 74%, 72%, and 71%, respectively. Longitudinal binomial regression revealed no significant differences at the  $p < .10$ , level in the percentage of completed assessments for any informant among comparison girls and boys and girls and boys with ADHD. There were not enough assessments in which only the parent or the youth participated to estimate differences, but a teacher did not participate in a substantial minority of the assessments. Although a sex difference was not found in the participation of teachers, when the response variable was based

on combined parent and teacher reports, the number of informants participating in each assessment was controlled.

The percentage of interviews of parents conducted by interviewers who were completely unaware of previous assessments in each successive follow-up assessment was 72%, 66%, 60%, 52%, 36%, 44%, and 59%, respectively. The percentage of blind interviews of youth in the four follow-up assessments in which youth were interviewed (Years 6, 7, 8, and 9) was 65%, 44%, 42%, and 43%, respectively. This issue was not germane to teacher reports, as different teachers completed questionnaires each year. Longitudinal binomial regression did not reveal a sex difference in blind interviews of parents,  $z = -0.52$ ,  $p = .60$ , or youth,  $z = -1.36$ ,  $p = .17$ , over the 9 years. Nonetheless, blindness of interviews was controlled to control for any extraneous differences due to the blindness or lack of blindness of the interviewer.

**Medication and psychosocial treatment.** The percentage of the 118 children with ADHD in Year 1 who were taking psychoactive medications during the 12 months preceding each assessment was 20%, 52%, 53%, 58%, 56%, 58%, 57%, and 56%, respectively. The corresponding percentages for comparison children were 1%, 1%, 2%, 4%, 6%, 5%, 5%, and 5%. Longitudinal binomial regression did not reveal a sex difference in prescription medication over years, controlling for the diagnosis of ADHD in Year 1,  $z = -1.26$ ,  $p = .21$ . The percentage of children with ADHD in Year 1 who received psychosocial treatment during the 12 months preceding each of the seven assessments was 52%, 62%, 59%, 59%, 61%, 59%, 59%, and 62%, respectively. The corresponding percentages for comparison children were 6%, 9%, 11%, 9%, 17%, 15%, 12%, and 9%. Longitudinal binomial regression did not reveal a sex difference in receipt of psychosocial treatment, controlling for the diagnosis of ADHD in Year 1,  $z = -1.48$ ,  $p = .14$ .

In addition, psychosocial treatment and psychoactive medication were tested as potential time-varying methodologic covariates (0 = no, 1 = yes) in all models to determine if they were associated with fewer symptoms and less impairment in each assessment. When the treatments were significant in models, they always reflected higher levels of symptoms and impairment during years in which treatment was received. We interpreted this as reflecting treatment-seeking when symptoms and impairment were higher and did not control treatment to avoid overcontrolling group differences.

**Table 1.** Characteristics in the Year 1 Assessment of Girls and Boys Who Met Full Criteria for ADHD in Year 1 and Comparison Children

	Comparison				ADHD			
	Girls <sup>a</sup>		Boys <sup>b</sup>		Girls <sup>c</sup>		Boys <sup>d</sup>	
	% M	SD	% M	SD	% M	SD	% M	SD
Race-ethnicity								
Non-Hispanic White	62.5		63.7		65.0		62.2	
African American	33.3		30.4		25.0		31.6	
Hispanic or Other Groups	4.2		5.9		10.0		6.1	
Age in years	5.1	0.8	5.2	0.8	5.4	0.7	5.2	0.7
Family income in Year 1 (US\$)	49,114	31,023	47,730	36,580	41,400	39,358	39,494	34,700
Mean intelligence	101.1	12.7	104.5	14.5	94.6	11.0	92.3	12.6
Inattention symptoms	0.4	0.7	0.7	1.3	7.3	1.6	7.3	2.0
Hyperactivity-impulsivity symptoms	0.9	1.1	1.3	1.6	6.8	2.1	7.8	1.5
Conduct symptoms (ODD + CD)	0.3	0.6	1.0	1.8	5.0	3.2	6.3	3.8
Internalizing symptoms	1.9	2.5	1.9	2.6	5.4	5.0	5.3	4.1

Note: Logged family incomes were compared, but mean incomes are tabled. ADHD = attention-deficit/hyperactivity disorder; mean intelligence = mean of Standard-Binet scores in Years 1 and 2; ODD = oppositional defiant disorder; CD = conduct disorder; ODD + CD = total number of *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM-IV*]) ODD and *DSM-IV* CD symptoms; internalizing symptoms = number of *DSM-III-R* symptoms of anxiety disorders, major depression, and dysthymia.

<sup>a</sup>  $n = 24$ .

<sup>b</sup>  $n = 102$ .

<sup>c</sup>  $n = 20$ .

<sup>d</sup>  $n = 98$ .

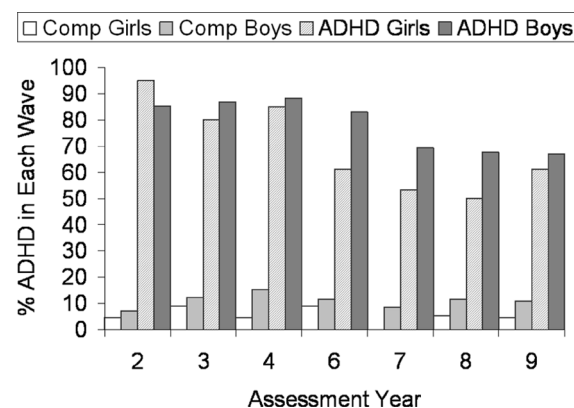
## Results

As shown in Table 1, there were no significant differences among the four groups in demographic characteristics, but both girls and boys who met criteria for ADHD in Year 1 had lower intelligence scores than comparison children, both  $p < .0001$ . Both girls and boys with ADHD in Year 1 exhibited greater numbers of symptoms of conduct problems and internalizing disorders in Year 1 than comparison children,  $p < .0001$ . There were no significant differences at  $p < .05$  between girls and boys who met criteria for ADHD on any of these Year 1 variables, however.

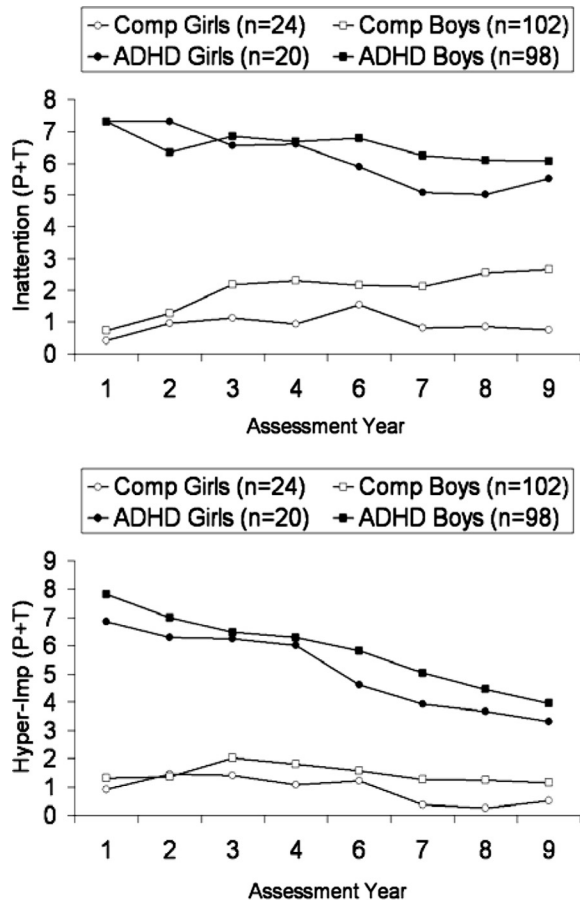
### Diagnosis and Symptoms of ADHD During Years 2 Through 9

Percentages of girls and boys in the ADHD and comparison groups meeting ADHD criteria each year are presented in Figure 1. Longitudinal binomial regression revealed that girls with ADHD in Year 1 were more likely than comparison girls to meet criteria for ADHD across Years 2 through 9,  $\beta = 3.92$ ,  $z = 6.81$ ,  $p < .0001$ . Boys with ADHD in Year 1 also were more likely than comparison boys to meet criteria for ADHD across Years 2 through 9,  $\beta = 3.58$ ,  $z = 11.39$ ,  $p < .0001$ . There was not a significant difference between girls and boys with ADHD in the proportion meeting criteria for ADHD across

Years 2 through 9,  $\beta = 0.29$ ,  $z = 0.98$ ,  $p = .32$ . To estimate effect sizes for the relative stability of the diagnosis of ADHD in girls and boys in easily understood terms, the proportions who met criteria for ADHD in either of the last two assessments (i.e., Year 8, Year 9, or both) were determined. A greater proportion of girls with ADHD in Year 1 (65.0%) than comparison girls (4.8%) met criteria for ADHD in Year 8, Year 9, or both, odds ratio (OR) = 31.24 (95%



**Figure 1.** The percentage of girls and boys who met *Diagnostic and Statistical Manual for Mental Disorders* (4th ed., American Psychiatric Association, 1994) criteria for attention-deficit/hyperactivity disorder (ADHD) in Year 1 and demographically matched comparison girls and boys who met *DSM-IV* criteria for ADHD in each annual assessment (no assessment in Year 5).



**Figure 2.** Mean numbers of *Diagnostic and Statistical Manual for Mental Disorders* (4th ed., American Psychiatric Association, 1994) inattention (left panel) and hyperactivity–impulsivity (right panel) symptoms reported by parents and teachers during each assessment among girls and boys who met *DSM-IV* criteria for attention-deficit/hyperactivity disorder (ADHD) in Year 1 and demographically matched comparison (Comp) girls and boys.

confidence interval (CI) = 3.1 to 316.0). Similarly, a greater proportion of boys with ADHD in Year 1 (77.2%) than comparison boys (14.7%) met criteria for ADHD 7 to 8 years later, OR = 23.42, 95% (CI = 10.3 to 53.3). Among children with ADHD in Year 1, boys were not more likely than girls to meet criteria for ADHD 7 to 8 years later (OR = 2.43, 95% CI = 0.7 to 8.5).

Similarly, as shown in the top panel of Figure 2, girls with ADHD in Year 1 were reported by parents and teachers to exhibit more inattention symptoms over Years 2 through 9 than comparison girls,  $\beta = 1.74$ ,  $z = 6.10$ ,  $p < .0001$ . The magnitude of this difference was quantified by comparing the mean number inattention symptoms during the last three assessments (Years 7 to 9) among girls with ADHD during Year 1 and comparison girls relative to the pooled standard deviation (Cohen, 1988). According to the

standards provided by Cohen, the size of the difference was large ( $d = 1.80$ ). Boys with ADHD in Year 1 similarly exhibited more inattention symptoms over time than comparison boys,  $\beta = 1.09$ ,  $z = 10.64$ ,  $p < .0001$ . The magnitude of the mean difference in inattention during Years 7 through 9 was large for boys ( $d = 1.46$ ).

The comparison between girls with and without ADHD in Year 1 did not interact significantly with time,  $\beta = -0.02$ ,  $z = -0.18$ ,  $p = .85$ , but the comparison between boys who did and did not meet criteria for ADHD in Year 1 did,  $\beta = -0.10$ ,  $z = -3.92$ ,  $p < .0001$ . This reflects a narrowing over successive years of the difference in inattention symptoms between boys with and without ADHD in Year 1. That is, inattention symptoms increased among comparison boys,  $\beta = 0.10$ ,  $z = 4.44$ ,  $p < .0001$ , but did not decline significantly among boys with ADHD,  $\beta = -0.00$ ,  $z = -0.33$ ,  $p = .74$ . Girls with ADHD did not exhibit higher levels of inattention across Years 2 through 9 than boys with ADHD,  $\beta = 0.04$ ,  $z = 0.60$ ,  $p = .55$ .

The lower panel of Figure 2 shows that girls with ADHD in Year 1 exhibited more hyperactivity–impulsivity symptoms over Years 2 through 9 than comparison girls,  $\beta = 1.57$ ,  $z = 6.43$ ,  $p < .0001$ . The magnitude of this difference across Years 7 through 9 was large ( $d = 1.82$ ). Boys with ADHD also exhibited more hyperactivity–impulsivity symptoms over time than comparison boys,  $\beta = 1.34$ ,  $z = 11.94$ ,  $p < .0001$ . The magnitude of this difference across Years 7 through 9 was large ( $d = 1.43$ ). Boys with ADHD in Year 1 did not exhibit significantly more hyperactivity–impulsivity symptoms during Years 2 through 9 than girls with ADHD in Year 1,  $\beta = 0.11$ ,  $z = 1.29$ ,  $p = .20$ .

### Functional Impairment During Years 2 Through 9

All analyses of group differences in impairment conservatively controlled family income, child intelligence, Year 1 internalizing (anxiety + depression) symptoms, and Year 1 conduct problems (oppositional defiant disorder + CD symptoms) in addition to the methodologic covariates (site, cohort, and the number of informants when relevant). As shown in the top panel of Figure 3, girls with ADHD in Year 1 received lower parent CGAS ratings during Years 2 through 9 than comparison girls,  $\beta = -15.04$ ,  $z = -5.58$ ,  $p < .0001$ , indicating less adaptive functioning. The magnitude of this difference over the last three assessments was large ( $d = 2.06$ ). Boys with ADHD also received lower parent CGAS ratings over Years 2 through 9 than comparison boys,  $\beta = -8.44$ ,  $z = -5.05$ ,  $p < .0001$ . The magnitude



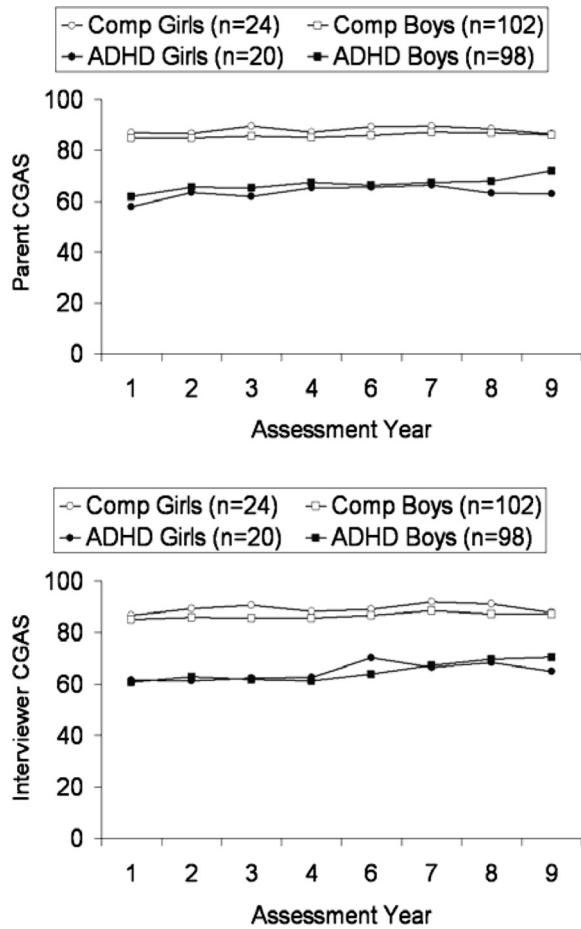


Figure 3. Mean ratings on the Children's Global Assessment Scale (CGAS) by parents (top panel) and interviewers (lower panel) during each assessment among girls and boys who met *Diagnostic and Statistical Manual for Mental Disorders* (4th ed., American Psychiatric Association, 1994) criteria for attention-deficit/hyperactivity disorder (ADHD) in Year 1 and demographically matched comparison (Comp) girls and boys.

of this difference over the last three assessments was large ( $d = 1.52$ ). Boys with ADHD received slightly but significantly higher (more favorable) parent CGAS ratings ( $M = 69.0$ ) over Years 2 through 9 than girls with ADHD ( $M = 64.1$ ),  $\beta = 5.79$ ,  $z = 2.39$ ,  $p < .02$  ( $d = .35$  for the mean of the last three assessments).

The lower panel of Figure 3 shows that girls with ADHD in Year 1 received lower interviewer CGAS ratings during Years 2 through 9 than comparison girls,  $\beta = -14.16$ ,  $z = -5.97$ ,  $p < .0001$ . The magnitude of difference averaged over the last three assessments was large ( $d = 2.08$ ). Boys with ADHD also received lower interviewer CGAS ratings over Years 2 through 9 than comparison boys,  $\beta = -9.92$ ,  $z = -6.85$ ,  $p < .0001$ . When the interaction with time was tested, it was significant for boys,  $\beta = 1.05$ ,  $z = 3.19$ ,  $p < .005$ ,

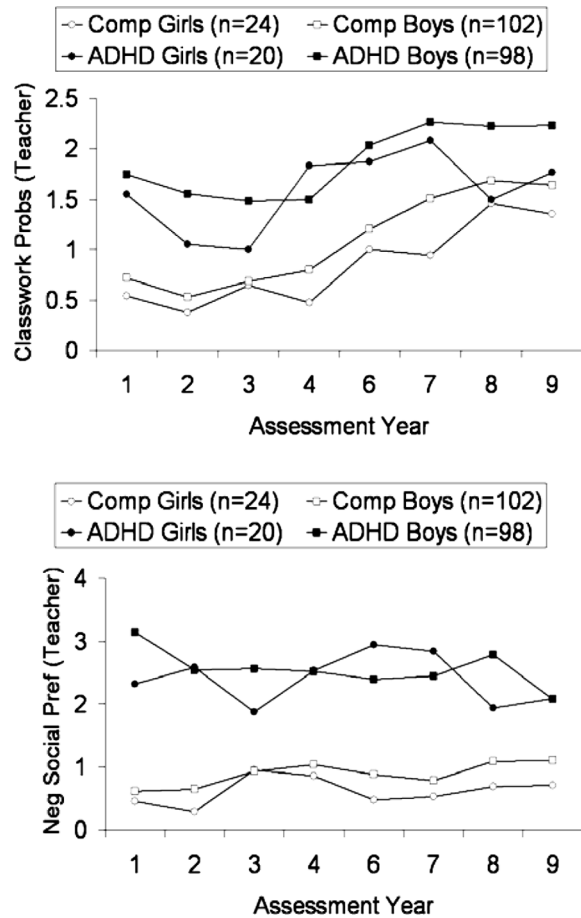


Figure 4. Mean teacher ratings of problems in the completion and accuracy of academic work in the classroom (top panel) and negative social preference scores (lower panel) during each assessment among girls and boys who met *Diagnostic and Statistical Manual for Mental Disorders* (4th ed., American Psychiatric Association, 1994) criteria for attention-deficit/hyperactivity disorder (ADHD) in Year 1 and demographically matched comparison (Comp) girls and boys. Higher negative social preference scores indicated that the child was rated as more disliked than liked by classmates.

indicating that interviewer CGAS ratings for boys with ADHD rose somewhat over time bringing them closer to those of comparison boys. Nonetheless, the difference in the means of interviewer CGAS ratings over the last three assessments between boys with ADHD and comparison boys was still large ( $d = 1.67$ ). Boys and girls with ADHD did not differ in their interviewer CGAS ratings,  $\beta = 3.13$ ,  $z = 1.41$ ,  $p = .16$ .

As shown in the top panel of Figure 4, teachers rated girls with ADHD as having more classroom academic problems across Years 2 through 9 than comparison girls,  $\beta = 0.42$ ,  $z = 3.03$ ,  $p < .005$ . When the interaction with time was tested, it was significant,  $\beta = -0.14$ ,  $z = -2.03$ ,  $p < .05$ ,

reflecting a greater increase in schoolwork problems in comparison girls than in girls with ADHD. The magnitude of the difference in schoolwork problems over the last three assessments between girls with and without ADHD was still large ( $d = 1.02$ ), but their means were nearly identical in Year 8. Similarly, boys with ADHD had more schoolwork problems during Years 2 through 9 than comparison boys,  $\beta = 0.30$ ,  $z = 3.28$ ,  $p < .005$ . When the interaction with time was tested it was significant,  $\beta = -0.11$ ,  $z = -4.20$ ,  $p < .0001$ , reflecting a greater increase in schoolwork problems in comparison boys than in boys with ADHD. The magnitude of this difference over the last three assessments was moderate ( $d = .74$ ). Boys and girls with ADHD did not differ significantly in schoolwork problems during Years 2 through 9,  $\beta = 0.12$ ,  $z = 1.40$ ,  $p = .16$ .

As shown in the lower panel of Figure 4, teachers gave girls with ADHD more negative social preference scores than comparison girls during Years 2 through 9,  $\beta = 1.05$ ,  $z = 3.56$ ,  $p < .0005$ . The magnitude of this difference over the last three assessments was large ( $d = 1.28$ ). Boys with ADHD similarly received higher negative social preference ratings than comparison boys,  $\beta = 0.74$ ,  $z = 4.53$ ,  $p < .0001$ . The magnitude of this difference also was large ( $d = 1.02$ ). There was not a significant difference in social preference ratings between girls and boys with ADHD,  $\beta = -0.06$ ,  $z = -.39$ ,  $p = .70$ .

By Year 9, 55.0% of girls and 60.2% of boys with ADHD were reported to have had at least one unintentional injury, compared with 25.0% of comparison girls and 29.4% of comparison boys. Logistic regression controlling age, site, and cohort revealed that girls with ADHD were more likely to have experienced an unintentional injury than comparison girls, OR = 3.86 (95% CI = 1.05 to 14.20). Similarly, boys with ADHD were more likely to have experienced an unintentional injury than comparison boys, OR = 3.54 (95% CI = 1.94 to 6.45). Boys with ADHD were not significantly more likely to have an unintentional injury than girls with ADHD, OR = 1.17 (95% CI = 0.44 to 3.12).

By Year 9, 65.0% of girls and 79.6% of boys with ADHD had been placed in a special education program at least once, compared with 12.5% of comparison girls and 31.4% of comparison boys. Logistic regression controlling age, site, and cohort revealed that girls with ADHD were more likely to have been placed in special education than comparison girls, OR = 15.28 (95% CI = 3.22 to 72.53). Similarly, boys with ADHD were significantly more likely to have been placed in special education than control boys, OR = 8.8

(95% CI = 4.5 to 17.3). Boys with ADHD were not significantly more likely to have been placed in special education than girls with ADHD, OR = 2.0 (95% CI = 0.6 to 5.7).

### Symptoms of Other Mental Disorders During Years 6 Through 9

To comprehensively describe the mental health outcomes of girls and boys with ADHD during Year 1, the groups were compared on their levels of symptoms of CD, depression, and anxiety disorders during early adolescence. Because youth are considered to be a valid informant on these disorders, analyses were based on data from the assessments in which both parent and youth were interviewed (i.e., CD and depression in Years 6 to 9 and anxiety disorders in Years 7 to 9). As shown in Figure 5, girls with ADHD in Year 1 exhibited more CD symptoms across Years 6 through 9 than comparison girls,  $\beta = 1.85$ ,  $z = 3.54$ ,  $p < .0005$ . The magnitude of this difference during the last three assessments was large ( $d = .86$ ). This group comparison was significant ( $\beta = 1.40$ ,  $z = 2.21$ ,  $p < .03$ ) even when conservatively controlling for three variables that were partly confounded with ADHD in Year 1: family income, the number of internalizing symptoms in Year 1, and the number of conduct problems in Year 1. Similarly, controlling for all methodologic covariates, boys with ADHD in Year 1 exhibited more CD symptoms across Years 6 through 9 than comparison boys,  $\beta = 1.64$ ,  $z = 6.10$ ,  $p < .0001$ .

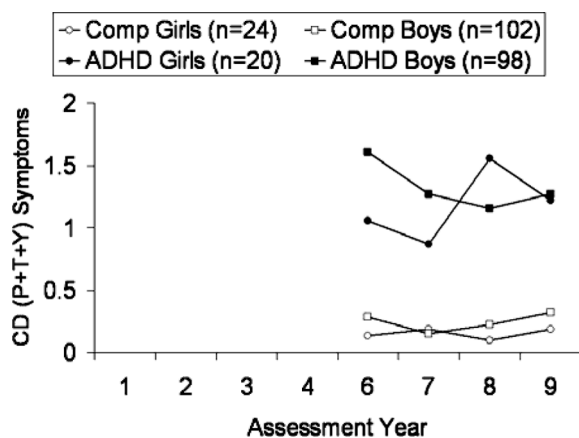


Figure 5. Mean numbers of conduct disorder (CD) symptoms reported by the parent, teacher, or youth during each assessment among girls and boys who met *Diagnostic and Statistical Manual for Mental Disorders* (4th ed., American Psychiatric Association, 1994) criteria for attention-deficit/hyperactivity disorder (ADHD) in Year 1 and demographically matched comparison (Comp) girls and boys.

The magnitude of this difference during the last three assessments was large ( $d = .92$ ). This group comparison was significant ( $\beta = 0.76$ ,  $z = 2.45$ ,  $p < .02$ ) even controlling for family income, the number of internalizing symptoms in Year 1, and the number of conduct problems in Year 1. Boys with ADHD did not exhibit significantly more CD symptoms during Years 6 through 9 than girls with ADHD,  $\beta = 0.13$ ,  $z = 0.42$ ,  $p < .001$ .

As shown in the top panel of Figure 6, girls with ADHD in Year 1 exhibited more symptoms of major depression across Years 6 through 9 than comparison girls,  $\beta = 0.92$ ,  $z = 4.76$ ,  $p < .0001$ . The magnitude of this difference during the last three assessments was large ( $d = 1.13$ ). This group comparison was significant ( $\beta = 0.63$ ,  $z = 3.06$ ,  $p < .005$ ) even controlling for family income, the number of internalizing symptoms in Year 1, and

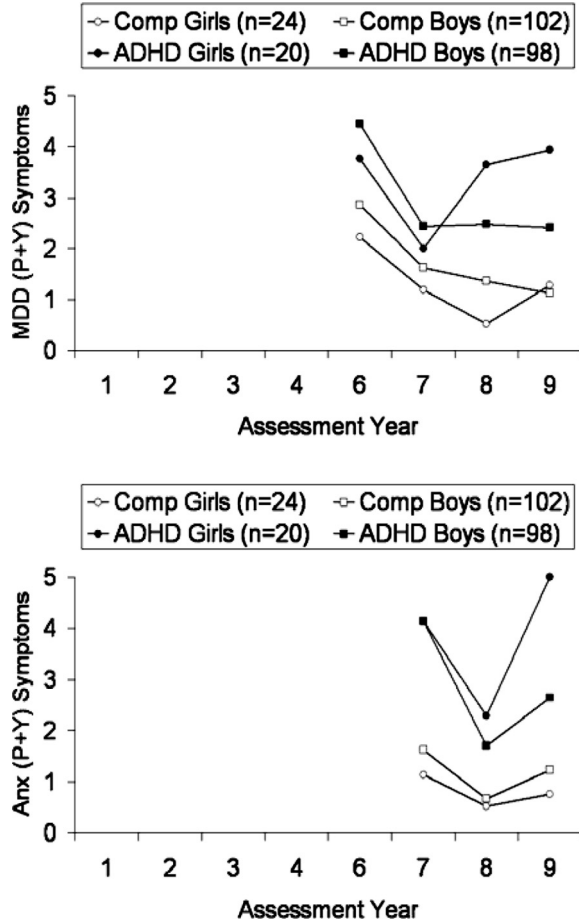


Figure 6. Mean numbers of symptoms of major depression (top panel) and anxiety disorders (lower panel) reported by the parent or youth during each assessment among girls and boys who met *Diagnostic and Statistical Manual for Mental Disorders* (4th ed., American Psychiatric Association, 1994) criteria for attention-deficit/hyperactivity disorder (ADHD) in Year 1 and demographically matched comparison (Comp) girls and boys.

the number of conduct problems in Year 1. Similarly, boys with ADHD in Year 1 exhibited more depression symptoms across Years 6 through 9 than comparison boys,  $\beta = 1.64$ ,  $z = 6.10$ ,  $p < .0001$ . The magnitude of this difference during the last three assessments was medium ( $d = .71$ ). This group comparison was significant ( $\beta = 0.29$ ,  $z = 2.53$ ,  $p < .02$ ) even controlling for family income, the number of internalizing symptoms in Year 1, and the number of conduct problems in Year 1. Girls with ADHD did not exhibit more depression symptoms across Years 6 through 9 than boys with ADHD,  $\beta = 0.18$ ,  $z = 1.47$ ,  $p = .14$ . However, when the interaction with time was tested, it was significant,  $\beta = -0.25$ ,  $z = -2.45$ ,  $p < .02$ . This interaction reflects slightly higher levels of depression among boys with ADHD than girls with ADHD during Years 6 and 7, and a steeper increase in depression symptoms into Years 8 and 9 (when the youth were 11–13 and 12–14 years old, respectively) among girls with ADHD than boys with ADHD.

As shown in the lower panel of Figure 6, girls with ADHD in Year 1 exhibited more symptoms of anxiety disorders across Years 7 through 9 than comparison girls,  $\beta = 1.58$ ,  $z = 4.34$ ,  $p < .0001$ . The magnitude of this difference over the last three assessments was large ( $d = .84$ ). This group comparison was significant ( $\beta = 0.98$ ,  $z = 2.78$ ,  $p < .01$ ) even controlling for family income, the number of internalizing symptoms in Year 1, and the number of conduct problems in Year 1. Similarly, boys with ADHD in Year 1 exhibited more anxiety symptoms across Years 7 through 9 than comparison boys,  $\beta = 0.91$ ,  $z = 5.52$ ,  $p < .0001$ . The magnitude of this difference over the last three assessments was moderate ( $d = .76$ ). This group comparison was significant ( $\beta = 0.42$ ,  $z = 2.02$ ,  $p < .05$ ) even controlling for family income, the number of internalizing symptoms in Year 1, and the number of conduct problems in Year 1. Girls with ADHD did not exhibit more anxiety symptoms across Years 7 through 9 than boys with ADHD,  $\beta = 0.37$ ,  $z = 1.50$ ,  $p = .14$ . When the interaction with time was tested, however, it was significant,  $\beta = -0.37$ ,  $z = -2.14$ ,  $p < .05$ , reflecting a steeper increase in anxiety symptoms from Year 8 to Year 9 among girls with ADHD than boys with ADHD.

## Discussion

These findings strongly suggest that the diagnosis of ADHD at 4 to 6 years of age exhibits predictive validity for both girls and boys. That is, the diagnosis of ADHD predicted continuing symptoms of

ADHD and multiple forms of related functional impairment as reported by three informants (i.e., parents, teachers, and interviewers) over 8 years in both sexes. These findings also indicate that ADHD in young children is a clinically significant disorder partly because it predicts high levels of symptoms of other mental disorders in early adolescence, even when Year 1 internalizing symptoms and conduct problems were controlled. That is, ADHD demonstrates both marked homotypic continuity (i.e., childhood ADHD predicts future ADHD) and heterotypic continuity (i.e., childhood ADHD predicts future symptoms of other mental disorders) in both sexes. Both girls and boys with ADHD exhibited higher levels of CD symptoms in early adolescence than comparison youth.

On the other hand, girls with ADHD in Year 1 not only exhibited more anxiety and depression symptoms during early adolescence than comparison girls, but they also showed significantly steeper increases in these symptoms than boys with ADHD during these years. This may be one reason why boys with ADHD received slightly but significantly more favorable parent CGAS ratings over time than girls with ADHD. The steeper increases in adolescent depression and anxiety among girls with ADHD could also explain why ADHD has been found to predict more inpatient psychiatric treatment in girls than boys (Dalsgaard et al., 2002). It will be possible to conduct future assessments of this sample when they have passed through adolescence to determine how many girls and boys exhibit high levels of impairing symptoms of these disorders and the numbers who met *DSM-IV* diagnostic criteria for major depression, serious anxiety disorders, and CD through adolescence. Any clinical implications of these findings should be interpreted with caution until these future assessments are completed, but it is important to note that both subthreshold and threshold numbers of symptoms of depression during adolescence predict serious adult mental disorders (Fergusson, Horwood, Ridder, & Beautrais, 2005). Thus, it seems appropriate to cautiously consider young children who meet criteria for ADHD to be at increased risk for later impairing problems of emotion, especially young girls with ADHD.

These findings did not support previous findings of lower levels of comorbid conduct problems in girls with ADHD than boys with ADHD (Abikoff et al., 2002; Gaub & Carlson, 1997; Gershon, 2002; Levy et al., 2005) but did confirm previous findings that ADHD is more strongly related to emotional problems in girls than boys. This is of great interest as it raises the possibility of differences in the fundamental nature of ADHD

in girls and boys, an important possibility that deserves careful further study.

The relatively small number of girls with ADHD was a limitation of this study, but the differences between girls with ADHD in Year 1 and comparison girls all reflected large effect sizes, and all tests of predictive validity for girls with ADHD were statistically significant. Indeed, the only moderate-sized difference between ADHD and comparison youth was found for boys. Even the lower-bound estimates for tests of predictive validity for girls were well in the range of statistical significance. For example, the 95% CI for the OR comparing girls with and without ADHD in Year 1 on the diagnosis of ADHD in Years 8 and 9 was 3.1 to 316.0. Even if the true effect size were half as large as the estimated lower bound, it would still be above 1.0. Thus, it seems unlikely that the relatively small number of girls meaningfully limited our power to detect large differences between girls with ADHD in Year 1 and comparison girls.

On the other hand, it is certainly possible that small differences between girls and boys with ADHD were not detected in this study due to limited power. That is, the diagnosis of ADHD appears to be valid for both girls and boys, but we cannot rule out the possibility that it is more strongly related to the indexes of predictive validity in one sex. Fortunately, the former issue is far more important for assessing the validity of diagnostic criteria among both girls and boys than the latter issue.

A potential bias in the question about the child's physical injuries should be noted. The question was worded to exclude injuries resulting from abuse or accidents caused by adults, but focusing the question on accidents caused by the child's carelessness or impulsivity could bias the parents of children with ADHD to report more such injuries because they view their children as careless and impulsive. However, other studies of unintentional injuries lend credence to these findings. Previous studies that used definitions of injury that were not biased in the same way had similar results, such as finding that ADHD predicts bone fracture injuries (Rowe, Maughn, & Goodman, 2004). Similarly, ADHD has been found to predict risky behavior in simulated vehicle driving tests (Barkley, Murphy, & Kwasnik, 1996) and simulated road crossing (Clancy, Rucklidge, & Owen, 2006). Thus, across studies, there seems to be little doubt that ADHD places children at high risk for physical injury.

The present findings are particularly important because concern has been raised that girls who are at risk for mental disorders in adolescence and adulthood may be underidentified relative to boys in childhood (Hartung & Widiger, 1998). If

future studies confirm the finding that the diagnosis of ADHD around the time of school entry predicts a broad range of important symptoms of mental disorders in girls during adolescence, it would mean that the diagnosis of ADHD is not only valid for girls but facilitates the early identification of girls who are at high risk for later serious mental disorders.

## References

- Abikoff, H. B., Jensen, P. S., Arnold, L. L. E., Hoza, B., Hechtman, L., Pollack, S., et al. (2002). Observed classroom behavior of children with ADHD: Relationship to gender and comorbidity. *Journal of Abnormal Child Psychology*, *30*, 349–359.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Barkley, R. A., Murphy, K. R., & Kwasnik, D. (1996). Motor vehicle driving competencies and risks in teens and young adults with attention deficit hyperactivity disorder. *Pediatrics*, *98*, 1089–1096.
- Barry, R. J., Clarke, A. R., McCarthy, R., & Selikowitz, M. (2006). Age and gender effects in EEG coherence: 3. Girls with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, *117*, 243–251.
- Baving, L., Laucht, M., & Schmidt, M. H. (1999). Atypical frontal brain activation in ADHD: Preschool and elementary school boys and girls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*, 1363–1371.
- Bird, H. R., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*, 78–85.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- 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.
- Cicchetti, D. & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*, 597–600.
- Clancy, T. A., Rucklidge, J. J., & Owen, D. (2006). Road-crossing safety in virtual reality: A comparison of adolescents with and without ADHD. *Journal of Clinical Child and Adolescent Psychology*, *35*, 205–215.
- Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology*, *18*, 557–570.
- Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, *60*, 837–844.
- Crick, N. R. & Zahn-Waxler, C. (2003). The development of psychopathology in females and males: Current progress and future challenges. *Development and Psychopathology*, *15*, 719–742.
- Dahmen, G., Rochon, J., König, I. R., & Ziegler, A. (2004). Sample size calculations for controlled clinical trials using generalized estimating equations (GEE). *Methods of Information in Medicine*, *43*, 451–456.
- Dalsgaard, S., Mortensen, P. B., Frydenberg, M., & Thomsen, P. H. (2002). Conduct problems, gender and adult psychiatric outcome of children with attention-deficit hyperactivity disorder. *British Journal of Psychiatry*, *181*, 416–421.
- Dishion, T. (1990). The peer context of troublesome child and adolescent behavior. In P. E. Leone (Ed.), *Understanding troubled and troubling youth* (pp. 128–153). Newbury Park, CA: Sage.
- Fabiano, G. A., Pelham, W. E., Waschbusch, D. A., Gnagy, E. M., Lahey, B. B., Chronis, A. M., et al. (2006). A practical measure of impairment: Psychometric properties of the Impairment Rating Scale in samples of children with attention-deficit/hyperactivity disorder and two school-based samples. *Journal of Clinical Child and Adolescent Psychology*, *35*, 369–385.
- Fergusson, D. M., Horwood, L. J., Ridder, E. M., & Beautrais, A. L. (2005). Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry*, *62*, 66–72.
- Gaub, M. & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 1036–1045.
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders*, *5*, 143–154.
- 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., & Hanson, K. S. (1994). Criterion validity of informants in the diagnosis of disruptive behavior disorders in children: A preliminary study. *Journal of Consulting & Clinical Psychology*, *62*, 410–414.
- 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.
- Hartung, C. M., Willcutt, E. G., Lahey, B. B., Pelham, W. E., Loney, J., Stein, M. A., et al. (2002). Gender differences in attention-deficit/hyperactivity disorder in preschool children. *Journal of Clinical Child and Adolescent Psychology*, *31*, 453–464.
- Hermens, D. F., Kohn, M. R., Clarke, S. D., Gordon, E., & Williams, L. M. (2005). Sex differences in adolescent ADHD: Findings from concurrent EEG and EDA. *Clinical Neurophysiology*, *116*, 1455–1463.
- Jensen, P. S., Rubio-Stipec, M., Canino, G., Bird, H. R., Dulcan, M. K., Schwab-Stone, M. E., et al. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry*, *38*, 1569–1579.
- 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., Applegate, B., Barkley, R. A., Garfinkel, B., McBurnett, K., Kerdyk, L., Greenhill, L., Hynd, G. W., Frick, P. J., Newcorn, J., Ollendick, T., Hart, E. L., Perez, D., Waldman, I., & Shaffer, D. (1994). DSM-IV field trials for oppositional defiant disorder and conduct disorder in children and adolescents. *American Journal of Psychiatry*, *151*, 1163–1171.

- Lahey, B. B., Pelham, W. E., Loney, J., Kipp, H., Ehrhardt, A., Lee, S. S., et al. (2004). Three-year predictive validity of *DSM-IV* attention deficit/hyperactivity disorder in children diagnosed at 4–6 years of age. *American Journal of Psychiatry*, *161*, 2014–2020.
- Lahey, B. B., Pelham, W. E., Loney, J., Lee, S. S., & Willcutt, E. (2005). Instability of the *DSM-IV* subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry*, *62*, 896–902.
- Lahey, B. B., Pelham, W. E., Stein, M. A., Loney, J., Trapani, C., Nugent, K., et al. (1998). Validity of *DSM-IV* attention-deficit/hyperactivity disorder for younger children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*, 695–702.
- Levy, F., Hay, D. A., Bennett, K. S., & McStephen, M. (2005). Gender differences in ADHD subtype comorbidity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*, 368–376.
- Mick, E., Biederman, J., Santangelo, S., & Wypij, D. (2003). The influence of gender in the familial association between ADHD and major depression. *Journal of Nervous and Mental Disease*, *191*, 699–705.
- 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.
- Newcorn, J. H., Halperin, J. M., Jensen, P. S., Abikoff, H. B., Arnold, L. L. E., Cantwell, D. P., et al. (2001). Symptom profiles in children with ADHD: Effects of comorbidity and gender. *Journal of American Academy Child & Adolescent Psychiatry*, *40*, 137–144.
- Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of *DSM-III-R* symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *31*, 210–218.
- Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology*, *20*, 51–63.
- 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.
- Rowe, R., Maughn, B., & Goodman, R. (2004). Childhood psychiatric disorder and unintentional injury: Findings from a national cohort study. *Journal of Pediatric Psychology*, *29*, 119–130.
- Sandstrom, M. J., & Cillessen, A. H. N. (2003). Sociometric status and children's peer experiences: Use of the daily diary method. *Merrill-Palmer Quarterly*, *49*, 427–452.
- Setterberg, S., Bird, H., & Gould, M. (1992). *Parent and interviewer version of the Children's Global Assessment Scale*. New York: New York State Psychiatric Institute.
- Shaffer, D., Fisher, P., Piacentini, J., Schwab-Stone, M., & Wicks, J. (1993). *Diagnostic interview schedule for children*. New York: Columbia University.
- Silberg, J. L., Rutter, M., Meyer, J., Maes, H., Hewitt, J., & Smonoff, E. (1996). Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *Journal of Child Psychology & Psychiatry*, *37*, 803–816.
- Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale*. Chicago: Riverside.
- Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, *42*, 121–130.

Received October 4, 2004

Accepted October 10, 2006