Biases in Ratings of Disruptive Behavior in Children: Effects of Sex and Negative Halos

Cynthia M. Hartung, Jill C. Van Pelt, Monica L. Armendariz and Laura A. Knight

J Atten Disord 2006; 9; 620

DOI: 10.1177/1087054705284233

Journal of Attention Disorders
Volume 9 Number 4
May 2006 620-630

10.1177/1087054705284233
http://jad.sagepub.com
hosted at
http://online.sagepub.com

Biases in Ratings of Disruptive Behavior in Children

Effects of Sex and Negative Halos

Cynthia M. Hartung
Jill C. Van Pelt
Monica L. Armendariz
Laura A. Knight
Oklahoma State University

Objective: Behavior disorders are more prevalent among boys than girls, but the etiology of this difference is unclear. Studies have not tested for sex bias in ratings as a contributing factor to the differential sex prevalence rates. However, there are several studies showing "negative halo effects" in ratings of boys (i.e., the presence of one type of behavior artificially inflating ratings of another behavior). These findings have only been extended to girls in one recent study. The current study is designed to test for sex difference in (a) ratings of boys and girls who exhibit the same degree of disruptive behavior and (b) negative halo effects. **Method:** Two hundred and thirty-nine college students participated. Sex differences in ratings are not found. Nonetheless, bidirectional negative halo effects are found for boys and girls (i.e., the presence of oppositionality artificially increased ratings of inattention and hyperactivity; the presence of inattention and hyperactivity artificially increased oppositionality). (*J. of Att. Dis. 2006; 9(4)620-630*)

Keywords: *behavior ratings; rater bias; sex bias; ADHD; children*

An abundance of studies have been conducted to better understand disruptive behavior disorders. Previous studies have suggested that the prevalence of ADHD is higher among males than females. Specifically, the male to female ratio is approximately 6:1 to 10:1 in clinical samples and 2:1 to 3:1 in nonreferred community samples (e.g., Barkley, 1998; Biederman et al., 2002; Gershon, 2002). In addition, Oppositional Defiant Disorder (ODD) appears to be more common in boys during childhood but equally common among girls and boys during adolescence (American Psychological Association [APA], 2000). A large majority of the studies of disruptive behavior disorders have not included adequate numbers of girls to allow for statistically meaningful comparisons between girls and boys (Gaub & Carlson, 1997; Hartung & Widiger, 1998). Therefore, the etiology of the higher prevalence of ADHD and ODD among boys is unclear. Without a better understanding of the etiology of this difference, it is virtually impossible to determine if *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text revision; *DSM-IV-TR*; APA, 2000) diagnostic criteria for ADHD and ODD are equally valid for males and females. The purpose of the present research was to examine the possible role of biases in ratings of children's behavior problems as a contributor to the differential sex prevalence rate in ADHD and ODD. Rater biases are one possible explanation for differential sex prevalence rates in ADHD and ODD. The rater bias theory, along with other theories, will be briefly reviewed.

Multiple hypotheses have been put forth in an attempt to understand the reasons for sex differences in these disorders (e.g., DeFries, 1989; Eme, 1992; James & Taylor, 1990; Keenan & Shaw, 1997; Rhee, Waldman, Hay, & Levy, 2001). These hypotheses can be grouped into two categories: (a) bias hypotheses and (b) true difference hypotheses. Bias hypotheses suggest that boys and girls

Authors' Note: This research was supported by start-up funds awarded to Dr. Hartung by Oklahoma State University. We would like to thank the following individuals for their contributions to this project: Jami Bartgis, Elizabeth Lefler, Lori Brinlee, Dea Burger, Kim Pittman, Andrea Siano, Lauretta Tulley, and Ashlee Wennberg. In addition, we are grateful to several colleagues for reviewing previous versions of this article: Jennifer Callahan, Shelia Kennison, and Melanie Page. Address correspondence to Cynthia M. Hartung, PhD, Department of Psychology, Oklahoma State University, 215 North Murray, Stillwater, OK 74078; e-mail: cynthia.hartung@okstate.edu.

display equal amounts of the same behaviors but are rated differently. For example, the rater bias hypothesis suggests that disruptive boys will be rated as more hyperactive than equally disruptive girls. In contrast, true difference hypotheses suggest that there are actual behavioral differences between boys and girls that will explain the differential sex prevalence rates. For example, the constitutional variability model suggests that boys with ADHD have more etiological variability than girls; therefore, girls show milder forms of ADHD (Eme, 1992; James & Taylor, 1990). The polygenetic multiple threshold model suggests that environmental and/or genetic factors combine to create a vulnerability for ADHD (DeFries, 1989; Rhee et al., 2001). Because girls have ADHD less frequently, this model suggests that girls have a higher threshold for ADHD and, therefore, need a greater liability to develop the disorder. Next, the differential socialization hypothesis suggests that girls are less likely to display ADHD symptoms because of socialization in spite of a similar predisposition to the disorder (Keenan & Shaw, 1997). Finally, the differential comorbidity hypothesis involves bias and true differences. This hypothesis suggests that boys are rated as more hyperactive and inattentive than girls because they are more likely to display comorbid oppositional and conduct disordered behaviors that draw more attention to their behavior in general.

Numerous studies have attempted to test these various hypotheses and have found some support for the differential comorbidity hypothesis, the polygenetic multiple threshold model, and the differential socialization hypothesis (e.g., DeFries, 1989; James & Taylor, 1990; Keenan & Shaw, 1997; Rhee et al., 2001). Nonetheless, the literature is largely inconclusive as to whether any or several of these hypotheses can account completely for the differential sex prevalence rates. In addition, no studies have specifically addressed the rater bias hypothesis, based on sex of child, as a possible contributing factor to the differential sex prevalence rates for behavior disorders.

Although no studies have specifically addressed sex bias in parent and teacher ratings of children's ADHD and/or ODD behaviors, another type of bias, negative halo effects, has been identified in rating scale measures of ADHD and ODD. Halo effects occur when a person's behavior in one domain is assumed to be representative of that person's behavior in other domains. Halo effects can be positive or negative. Positive halo effects occur when desirable behavior in one area leads one to assume that an individual will display similarly desirable behaviors in other areas. Negative halo effects occur when undesirable behaviors in one area lead to assumptions of undesirable behaviors in other domains. With regard to previous studies of negative halo effects, researchers have found the presence of oppositionality to increase the likelihood of inattentive or hyperactive ratings even in the absence of these symptoms (Abikoff, Courtney, Pelham, & Koplewicz, 1993; Schachar, Sandberg, & Rutter, 1986; Stevens, Quittner, & Abikoff, 1998). Although these studies showed negative halo effects for boys, they did not examine whether these effects are generalizable to girls. A more recent study examined sex differences in halo effects (Jackson & King, 2004). A review of these studies of negative halo effects follows.

Schachar et al. (1986) found that teachers rated boys who displayed defiance but not hyperactivity as defiant and hyperactive. This was an unidirectional finding where boys who displayed hyperactive behavior only were not rated as having symptoms of oppositionality. Stated differently, the presence of oppositionality artificially inflated levels of hyperactivity; however, the presence of hyperactivity did not affect levels of oppositionality. Similarly, Abikoff et al. (1993) found that teachers who watched a videotape of a boy who displayed oppositionality only rated the boy as having symptoms of oppositionality plus inattention and overactivity. Again, this finding was unidirectional. A 1998 study by Stevens, Quittner, and Abikoff revealed similar results. In this study, teachers who watched a videotape of a boy with oppositionality only had a tendency to rate the boy as inattentive and hyperactive in addition to rating him as oppositional. However, they did not rate a boy with inattention and hyperactivity symptoms as displaying oppositionality symptoms. All of these studies showed unidirectional negative halo effects in teachers' ratings of boys' behavior such that the presence of oppositionality artificially inflated ratings of ADHD (i.e., inattention and hyperactivity). However, the behavior of girls was not rated in any of these studies.

Recently, Jackson and King (2004) conducted a study designed to examine sex differences in halo effects for teacher ratings of disruptive behavior. Jackson and King created videotapes using scripts from the Abikoff et al. (1993) study. The Abikoff study included three tapes of boys (i.e., normal boy, ADHD only, and ODD only). Jackson and King used child actors to create male and female versions of each of these tapes. All teachers viewed a tape of a typical child followed by either an ADHD-only child or an ODD-only child of the same sex. Results indicated bidirectional negative halo effects such that the presence of oppositionality artificially inflated ratings of inattention and hyperactivity and the presence of inattention and hyperactivity artificially inflated ratings of oppositionality. There were no main effects of sex.

Therefore, the negative halo effect finding was extended to girls. Nonetheless, there were interactions with sex such that the presence of oppositionality resulted in significantly higher ratings of hyperactivity in the boy than in the girl. Conversely, the presence of inattention and hyperactivity resulted in significantly higher ratings of oppositionality in the girl than in the boy. This study suggests that negative halo effects may bias teacher reports of disruptive behavior for boys and girls. In addition, the tendency to assume that a child with ADHD has symptoms of ODD may be magnified for girls, whereas the tendency to assume that a child with ODD has symptoms of ADHD may be magnified for boys. Therefore, differential effects of sex on negative halos could result in more boys being incorrectly diagnosed with ADHD than girls and more girls being incorrectly diagnosed with ODD than boys. Although Jackson and King have extended the negative halo effects findings to girls, they did not specifically examine sex biases in teacher ratings of disruptive behavior but only sex differences in negative halo effects. Given that Jackson and King were the first to include girls in a study of sex differences in disruptive behavior disorder ratings, additional studies are needed to determine if their extension of the negative halo effect to girls will be replicated. In addition, future studies with boys and girls should also examine sex differences in ratings of ADHD and ODD for children who are displaying similar levels of symptomatology.

Although sex bias in teacher ratings of children's behaviors were not directly examined, two additional studies found sex differences in teacher ratings of treatment acceptability and teacher referrals for ADHD evaluations. First, Pisecco, Huzinee, and Curtis (2001) examined teacher ratings of treatment acceptability for children with ADHD. In this study, 159 elementary schools teachers viewed videotapes depicting children with ADHD. Each teacher read one of six vignettes that differed based on sex of child and ADHD subtypes. Teachers were also asked to read descriptions of various interventions for ADHD, including behavioral and pharmacological treatments, and rate the acceptability and effectiveness of each treatment for the child in the vignette. Pisecco et al. (2001) found a main effect for type of treatment in that teachers preferred one of the behavioral treatments (i.e., daily home report card) above all others. In addition, a Type of Treatment × Sex of Child interaction on teachers ratings of acceptability was found. Specifically, they found that teachers were more strongly opposed to medication as a treatment for girls than for boys.

Second, a recent study by Sciutto, Nolfi, and Bluhm (2004) examined the possibility of a referral bias, based on sex of child, by teachers. These researchers asked 200 elementary school teachers to read a description of a child with behavioral problems. Teachers were asked to report the degree to which they would find this child's behavior disruptive in the classroom and the likelihood that they would refer the hypothetical child for a psychological evaluation. Teachers were randomly assigned to one of six between-subjects conditions. The conditions varied based on the sex of the child and the type of behaviors exhibited (i.e., inattention only, hyperactivity only, and hyperactivity plus aggression). Results indicated no main effect of sex of child, or Sex of Child × Type of Behaviors Interaction, on ratings of classroom disruption. However, the likelihood that teachers would refer the child for a psychological evaluation was significantly greater if the child was a boy than if the child was a girl. There was also a significant Sex of Child × Type of Behavior interaction such that the tendency to refer boys more often than girls was most pronounced in the hyperactivity only condition. Taken together, these two studies provide additional support for the possibility that sex biases in ratings of children's behaviors may exist.

The current study was designed to test for sex bias in ratings of children with disruptive behavior disorders in addition to examining sex differences in negative halo effects. Specifically, to test for sex bias, ratings of boys and girls who were described as exhibiting the same degree of disruptive behavior disorder symptomatology were examined. In addition, the presence of negative halo effects and possible sex differences in these effects were examined.

Several a priori hypotheses were tested. First, assuming the rater bias hypothesis is accounting for a portion of the differential sex prevalence rates in ADHD and ODD, it was expected that disruptive behavior disorder ratings for boys would be inflated compared to girls who were described as displaying the same degree of disruptive behavior. In addition, it was expected that negative halo effects would be found across disruptive behavior disorders. Specifically, it was expected that the presence of oppositionality would artificially inflate levels of inattention and hyperactivity. Similarly, it was expected that the presence of both inattention and hyperactivity would artificially inflate levels of oppositionality. Based on the results of Jackson and King (2004), it was also expected that these halo effects would be found for boys and girls.

Method

Participants

Two hundred and thirty-nine undergraduate college students completed the study (124 male and 115 female).

Participants' mean age was 20.19 (*SD* = 4.29). The ethnic composition of the sample based on participant self-report was 85.4% Caucasian, 4.6% Native American, 3.8% Asian, 2.5% African American, 1.3% Hispanic, and 2.5% Other. Participants were recruited via verbal solicitation from undergraduate psychology courses at a large state university. Participants were given extra credit for a psychology course in which they were enrolled in exchange for participating in the study.

Vignettes

Eight vignettes describing child behavior were developed for the present study. The vignettes were adapted from abnormal child psychology textbooks and clinical child psychology case presentation texts (Barkley, 2000; Oltmanns, Neale, & Davidson, 1999). Each vignette described either one emotional and/or behavioral disorder of childhood, including ADHD–Predominately Inattentive Type (ADHD-PI), ADHD Combined Type (ADHD-CT), ODD, and Major Depressive Disorder (MDD). The MDD vignette was included so that it could be determined whether any sex biases would be specific to externalizing disorders or possibly generalizable to internalizing disorders. In addition, the MDD vignette served as a comparison condition when examining the presence of negative halo effects across externalizing vignettes. Each vignette had two versions, with one having a boy as the main character and the other having a girl. Except for changes in the main character's first name and associated pronouns, both versions of the vignette for each disorder were identical. Vignettes were edited such that they were all approximately 350 words. Given that there were four disorders described and two sex versions for each disorder, there were a total of eight vignettes.

Measures

Demographics Form

A demographic form included questions about participants' sex; age; ethnicity; number of years of education completed; current living situation; occupation; household income; marital status; number of siblings; history of special education or learning disability; symptoms of emotional or behavioral disturbances in self, siblings, and parents; and symptoms of emotional or behavioral disturbances in children (if applicable).

Behavior Rating Scale

Vignette version. A 43-item behavior rating scale was developed for the present study using symptoms for specific disorders as presented in the *DSM-IV-TR*. The behavior rating scale included the nine symptoms of inattention and the nine symptoms of hyperactivity that define ADHD, the eight symptoms of ODD, and nine symptoms of MDD. This measure was very similar to the *DSM-IV* Disruptive Behavior Disorder Checklist (Pelham, Gnagy, Greenslade, & Milich, 1992), but MDD symptoms were also included. After reading and hearing a vignette describing a child's behavior, participants indicated how often the child was likely to display each symptom listed on the behavior rating scale using a 4-point scale that included *never/don't know*, *sometimes*, *often*, or *very often*.

Self-report childhood version. The same 43 items described above were included in this version of the rating scale. Participants were asked to indicate how often they experienced these symptoms during childhood using the same 4-point scale described above.

Self-report current version. Using the same 43-item scale, participants were asked to indicate how often they currently experience each of the symptoms.

Procedure

Participants completed the study in groups of four or fewer. After providing informed consent, participants were given a written copy of the first vignette. Participants were asked to read along while an audio recording of the vignette was played. Next, participants rated the child in the vignette using the behavior rating scale. After completing the rating scale, the second vignette was distributed, and the corresponding audio recording was played. Participants then rated the child in the second vignette. This procedure was repeated for the remaining vignettes. Participants were not allowed to discuss the vignettes, and they were seated separately to insure that rating scales were completed independently.

Each participant rated a vignette for each of the four disorders but was randomly assigned to rate two vignettes depicting a boy and two vignettes depicting a girl. The presentation of the vignettes in each condition was such that the sex of the child in the vignette was always different than the sex of the child in the vignette that preceded it. Counterbalancing also insured that equal numbers of men and women rated the boy and girl versions of each disorder.

After rating all four vignettes, participants completed the demographics form, the self-report childhood version of the behavior rating scale, and the self-report current version of the behavior rating scale. Finally, the partici-

pants were debriefed and thanked for their participation. Procedures for recruiting participants and study method were approved by the institutional review board at the university where the data were collected and these procedures were in compliance with the ethical standards of the APA.

Data Analyses

The total number of symptoms endorsed as occurring often or very often (e.g., total symptom count) on each behavioral dimension (i.e., inattention, hyperactivity, oppositionality, and depression) for each vignette (i.e., ADHD-PI, ADHD-CT, ODD, and MDD) served as dependent variables. Therefore, there were a total of 16 dependent variables.

Negative halo effects were expected to occur within externalizing behaviors (e.g., presence of oppositionality artificially inflating inattention and hyperactivity ratings) but not across externalizing and internalizing dimensions (e.g., presence of oppositionality artificially inflating depression). Thus, the MDD vignette served as a psychiatric comparison for determining the presence of negative halo effects among disruptive behavior disorders. To test for the presence of negative halo effects, a series of paired sample *t* tests were conducted by comparing nontarget externalizing dimensions to the nontarget internalizing dimension of depression.

Results

To examine our hypotheses, multiple statistical comparisons were conducted. Bonferroni corrections were calculated for each set of analyses (e.g., three sets of preliminary analyses, sex bias analyses, and negative halo analyses). Resulting alpha values are noted for each set of results.

Preliminary Analyses

Effects of participant symptomatology. A series of correlational analyses were conducted to determine whether participants' own levels of childhood or current symptomatology affected their perceptions of the target child's behavior. Specifically, ratings on each of the four behavioral dimensions from each of the four vignettes were correlated with relevant dimensions of participants' childhood and current ratings. For example, depression ratings resulting from each of the four vignettes were correlated with participants' childhood ratings of depression and participants' current ratings of depression. This procedure was repeated for inattention, hyperactivity, and oppositionality. For each symptom dimension, 8 correlations were conducted, for a total of 32 analyses. The Bonferroni correction resulted in an alpha level of .002. None of the correlations between self-report symptom dimensions and target symptom dimensions were significant at this level.

Effects of symptomatology in children of participants. Participants also reported on their parental status. If they indicated that they had children, they were asked to report symptomatology for their own children. To determine whether symptomatology in their own children affected their perceptions of the target child's behavior, additional correlational analyses were planned. However, only two participants reported having children of their own. Because of insufficient power for comparing parents to nonparents, these analyses were not conducted.

Effects of sex of rater. To examine possible effects of sex of rater, a series of 16 independent sample *t* tests were conducted for target and nontarget dimensions. The Bonferroni correction resulted in an alpha level of .003. None of the *t* tests were significant at this level; therefore, it was not necessary to control for sex of rater when examining possible differences based on sex of child.

Sex Bias Analyses

To examine possible sex biases based on sex of child, a series of independent sample *t* tests were conducted for target and nontarget dimensions. Contrary to expectations, none of the 16 *t* tests resulted in any significant sex differences. Results are shown in Table 1.

Negative Halo Analyses

To test for the presence of negative halo effects, a series of paired sample *t* tests were conducted by comparing nontarget disruptive behavior dimensions to the nontarget depression dimension. Therefore, a total of five comparisons were made. The Bonferroni correction resulted in an alpha level of .01. For inattention, negative halo effects were expected for the ODD vignette but not for the MDD vignette, when inattention was a nontarget dimension (see Figure 1). Consistent with this expectation, *t* tests showed that ratings of inattention were significantly higher for the ODD than for the MDD vignette, $t(238) = 9.84$, $p = .001$. The independent sample *t* test that was previously conducted did not reveal a significant sex difference on level of inattention for the ODD vignette (see Table 1).

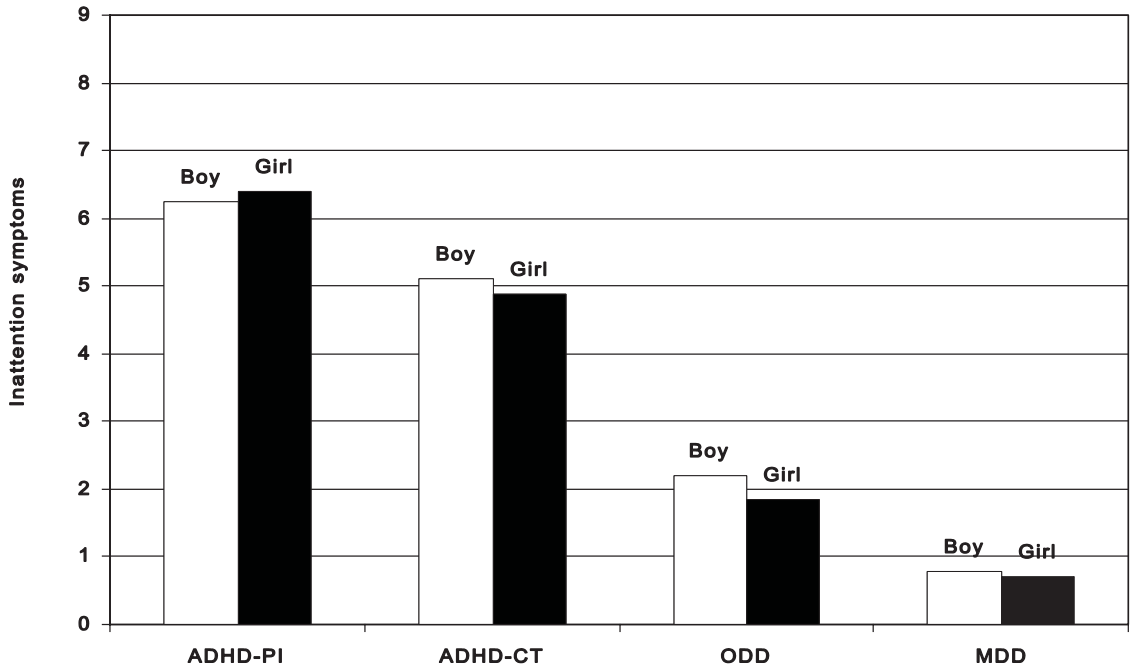
For hyperactivity, negative halo effects were expected for the ADHD-PI and ODD vignettes but not for the

Table 1
Means, Standard Deviations, and *t* Tests for Ratings of Girls and Boys on Target and Nontarget Dimensions

	Boys			Girls			*t* Test		
	M	*SD*	*n*	*M*	*SD*	*n*	*t*	*p*[a]	*df*
ADHD-PI vignette			123			116			237
Inattention[b]	6.26	1.72		6.38	1.70		0.54	.59	
Hyperactivity	2.53	2.29		2.18	2.04		1.24	.22	
Oppositionality	0.20	0.87		0.16	0.44		0.44	.66	
Depression	0.11	0.39		0.25	0.64		1.99	.05	
ADHD-CT vignette			117			122			237
Inattention[b]	5.10	2.05		4.89	1.93		0.81	.42	
Hyperactivity[b]	6.41	1.99		6.32	1.75		0.36	.72	
Oppositionality	3.21	2.22		3.39	2.22		0.61	.54	
Depression	0.76	0.93		0.64	0.83		1.11	.27	
ODD vignette			117			122			237
Inattention	2.21	2.32		1.84	2.07		1.30	.20	
Hyperactivity	3.21	2.57		3.12	2.58		0.25	.81	
Oppositionality[b]	7.05	1.55		7.07	1.49		0.11	.91	
Depression	1.06	1.14		1.02	0.09		0.25	.80	
MDD vignette			117			122			237
Inattention	0.79	1.28		0.70	1.20		0.59	.56	
Hyperactivity	0.09	0.35		0.12	0.61		0.42	.68	
Oppositionality	0.42	0.94		0.44	1.13		0.12	.90	
Depression[b]	6.63	1.26		6.32	1.48		1.70	.09	

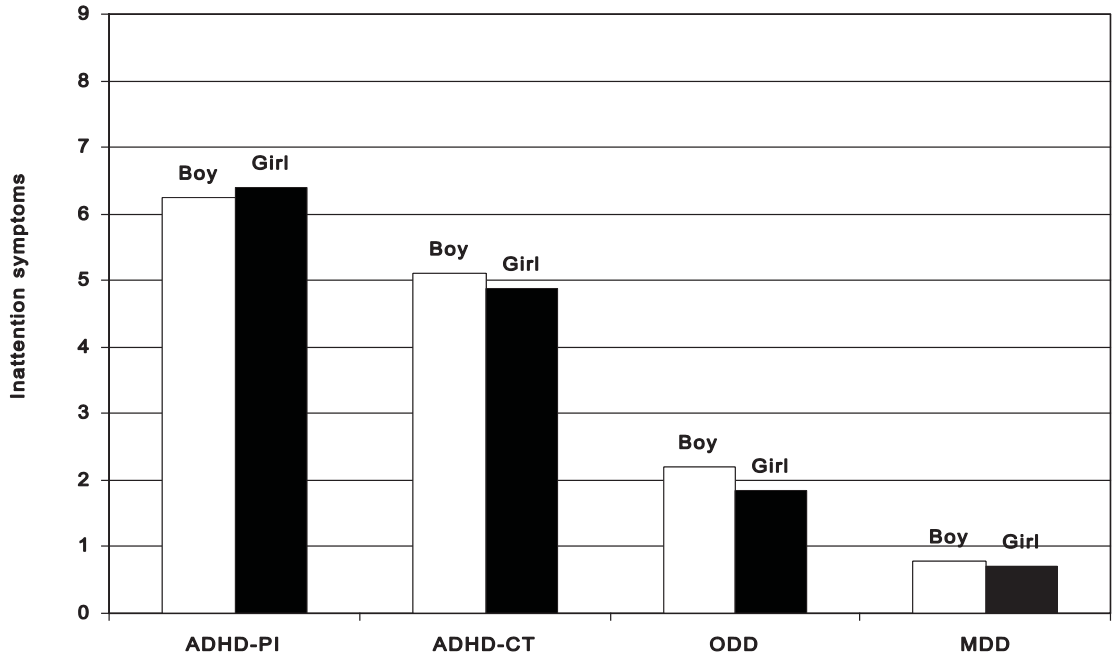
Note: PI = predominately inattentive; CT = combined type; ODD = Oppositional Defiant Disorder; MDD = Major Depressive Disorder.
a. Alpha value for this set of analyses was set at $p < .01$.
b. Target dimensions are the behaviors that were exhibited by the child in the vignette.

Figure 1
Inattention Ratings for Boys and Girls Across Vignettes



Note: PI = predominately inattentive; CT = combined type; ODD = Oppositional Defiant Disorder; MDD = Major Depressive Disorder.

Figure 2
Hyperactivity/Impulsivity Ratings for Boys and Girls Across Vignettes



Note: PI = predominately inattentive; CT = combined type; ODD = Oppositional Defiant Disorder; MDD = Major Depressive Disorder.

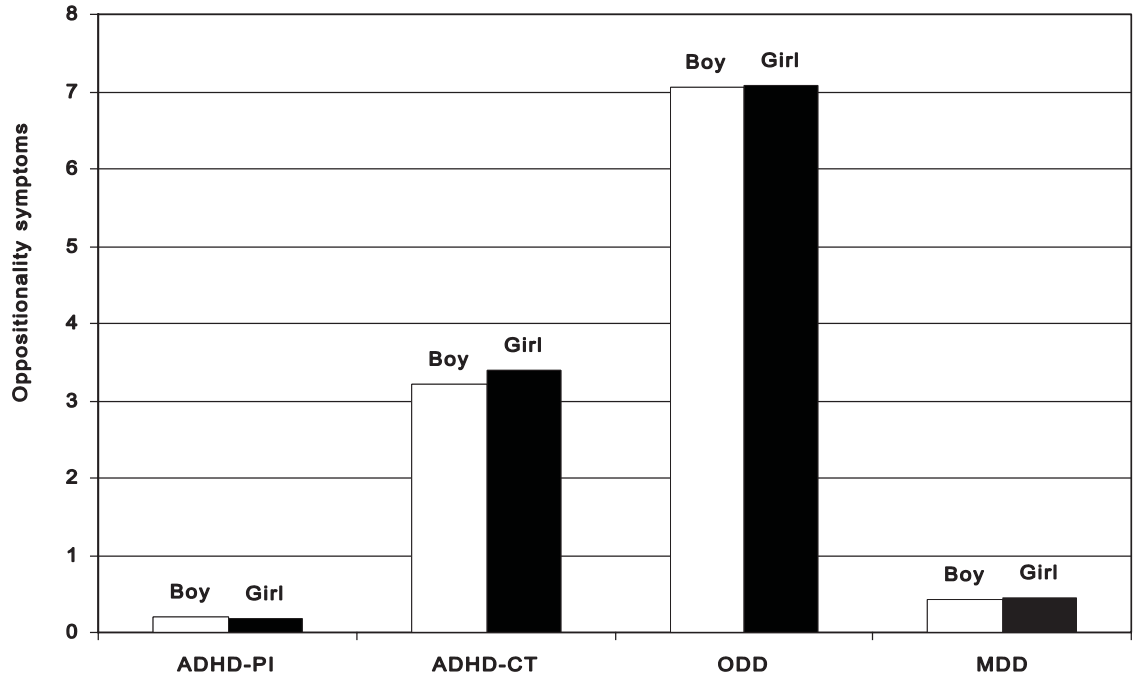
MDD vignette (see Figure 2). As expected, ratings of hyperactivity were significantly higher for the ADHD-PI than for the MDD vignette, $t(238) = 16.47$, $p = .001$. Similarly, ratings of hyperactivity were significantly higher for the ODD than for the MDD vignette, $t(238) = 18.31$, $p = .001$. The independent sample t tests that were previously conducted did not reveal a significant sex difference on levels of hyperactivity for the ADHD-PI or ODD vignettes (see Table 1).

For oppositionality, negative halo effects were expected for the ADHD-PI and ADHD-CT vignettes but not for the MDD vignette (see Figure 3). As expected, ratings of oppositionality were significantly higher for the ADHD-CT than for the MDD vignette, $t(238) = 20.16$, $p = .001$. In addition, ratings of oppositionality were significantly higher for the ADHD-PI than for the MDD vignette, $t(238) = 3.43$, $p = .001$. The independent sample t test that was previously conducted did not reveal a significant sex difference on level of oppositionality for the ADHD-CT or ADHD-PI vignettes (see Table 1).

Discussion

The purposes of the present study were to evaluate the rater bias hypothesis and negative halo effects as possible contributors to differential sex prevalence rates in ADHD and ODD. To evaluate the rater bias hypothesis, participants were presented with a series of vignettes depicting a boy or girl with symptoms of a childhood emotional or behavior disorder (i.e., ADHD-CT, ADHD-PI, ODD, or MDD). Results revealed that boys and girls were not rated differently on the target or nontarget dimensions depicted in the vignettes. These results do not support the rater bias hypothesis as contributing to the differential sex prevalence rates in ADHD and ODD. However, these findings must be qualified given that the study tested for sex bias in ratings of boys and girls who were highly symptomatic. In fact, the vignettes used in the current study depicted relatively severe cases of each disorder. Therefore, the current results specifically suggest that sex bias in behavior ratings does not account for differential diagnostic rates of these disorders in highly symptomatic children. This lack of ambiguity may also have decreased the probability that a sex bias would be observed. If the vignettes displayed children who were less severely symptomatic, there may have been a greater opportunity to observe sex differences. Consequently, the possibility of the existence of sex bias in behavior ratings of children who are less severely symptomatic has not yet been ruled out. Stated differently, it is possible that boys who are displaying a

Figure 3
Oppositionality Ratings for Boys and Girls Across Vignettes



Note: PI = predominately inattentive; CT = combined type; ODD = Oppositional Defiant Disorder; MDD = Major Depressive Disorder.

level of symptomatology that is closer to the threshold for the disorder might be diagnosed more often than girls who are showing the same level of symptomatology.

As expected, negative halo effects were found across certain behavioral dimensions. Results provided evidence of a bidirectional negative halo effect such that children with symptoms on one or more disruptive behavior dimensions were assumed to have symptoms on additional dimensions in spite of the lack of behavioral evidence to support this assumption. For inattention, the child in the ODD vignette was rated as having significantly higher levels of inattention than the child in the MDD vignette, which served as a comparison (see Figure 1). There was no sex difference in the magnitude of this halo effect. That is, the halo effect was present to a similar degree for boys and girls. This halo effect is consistent with previous studies (Abikoff et al., 1993; Schachar et al., 1986; Stevens et al., 1998). In addition, the fact that the halo effect generalizes to girls is consistent with the recent study by Jackson and King (2004). However, it should be noted that the level of inattention reported for the ODD vignette in the current study did not approach the clinical threshold. Specifically, the mean level of inattention in the ODD vignette was approximately two symptoms. Since six symptoms of inattention are necessary for a diagnosis of ADHD, this level does not suggest clinically significant problems with inattention. Thus, this negative halo effect does not appear to be severe enough to result in an ADHD-PI diagnosis for a child with ODD who is showing no symptoms of inattention. On the other hand, for a child with ODD who is displaying some symptoms of inattention, a negative halo effect could artificially elevate the inattention score to the clinically significant level.

For hyperactivity, the child in the ODD vignette and the child in the ADHD-PI vignette were rated as having significantly higher levels of hyperactivity than the child in the MDD vignette (see Figure 2). Again, there was no sex difference in the magnitude of this halo effect. This halo effect is also consistent with previous studies (Abikoff et al., 1993; Schachar et al., 1986; Stevens et al., 1998), and the generalization of this effect to girls is consistent with Jackson and King (2004). Similar to the halo effect on inattention, the levels of hyperactivity for the children in the ODD and ADHD-PI vignettes did not approach the clinical threshold (i.e., six symptoms). Specifically, the mean level of hyperactivity in the ODD vignette was approximately three symptoms, and the mean level in the ADHD-PI vignette was approximately two symptoms. Again, this negative halo effect is unlikely to result in a diagnosis of ADHD in a child with ODD who is showing no symptoms of hyperactivity. Similarly, this effect is unlikely to result in a diagnosis of ADHD-CT in a

child with ADHD-PI who is showing no symptoms of hyperactivity. However, in practice, there is a great degree of overlap in these symptoms, and a negative halo effect, for a child who is subthreshold on hyperactivity, could result in an incorrect diagnosis.

For oppositionality, the child in the ADHD-CT vignette was rated as having significantly higher levels of oppositionality than the child in the MDD vignette (see Figure 3). Again, there was no sex difference in the magnitude of this halo effect. This halo effect is also consistent with previous studies (Abikoff et al., 1993; Schachar et al., 1986; Stevens et al., 1998), and the generalization of this effect to girls is consistent with Jackson and King (2004). This halo effect was somewhat stronger than previous effects because levels of oppositionality did approach the clinical threshold (i.e., four symptoms). Specifically, the mean level of oppositionality in the ADHD-CT vignette was higher than three. Therefore, the child in the ADHD-CT vignette was rated almost high enough on oppositionality to obtain a clinical diagnosis of ODD in spite of the lack of behavior evidence to suggest that this child was oppositional. Therefore, rates of ODD in children with ADHD-CT might be artificially inflated.

In contrast to the ADHD-CT vignette, the child in the ADHD-PI vignette was not rated as having significantly higher levels of oppositionality than the child in the MDD vignette. This finding is not inconsistent with previous studies given that the majority did not find an ADHD-to-ODD negative halo effect but only a unidirectional effect from ODD to ADHD (Abikoff et al., 1993; Schachar et al., 1986; Stevens et al., 1998). Although Jackson and King (2004) found bidirectional negative halo effects, their conditions only included ADHD-CT and ODD but not ADHD-PI. Although a hyperactive-only condition was not included in the current study, these findings suggest that it may be the presence of hyperactivity that artificially inflates ratings of oppositionality and not the presence of inattention. Stated differently, the artificial inflation of oppositionality symptoms only occurred in the ADHD-CT vignette that described a child with both inattention and hyperactivity but not in the ADHD-PI vignette that described a child with inattention only. Given that a hyperactive-only vignette was not included, we cannot be sure whether the artificial inflation of oppositionality was caused by the presence of hyperactivity or the combined presence of inattention and hyperactivity.

Taken together, several statistically significant negative halo effects were identified. In all cases, the effects were found for both girls and boys, and there were no significant differences in the magnitude of these negative halo effects. Specifically, it was found that (a) oppositionality artificially inflated inattention and hyperactivity, (b) inattention artificially inflated hyperactivity, and (c) hyperactivity artificially inflated oppositionality. This study replicates the bidirectional negative halo effects found by Jackson and King (2004) and extends their finding to raters who are not teachers. In addition, as was the case with Jackson and King, the finding was evident for boys and girls. Taken together, the two studies suggest that the presence of ADHD-CT may artificially inflate rates of ODD for boys and girls. Similarly, the presence of ODD may artificially inflate rates of ADHD for boys and girls. The fact that there were no significant sex differences in the magnitude of these negative halo effects does not appear to support the hypothesis that sex differences in negative halo effects might partially account for the differential sex prevalence rates in ADHD and ODD. However, it is possible that there is an exponential impact of these negative halo effects on prevalence rates. Given that these negative halo effects might increase rates of ADHD and ODD in boys and girls but that ADHD and ODD are thought to be more prevalent in boys than girls, these halo effects might have a relatively greater impact on prevalence estimates for boys than for girls. Therefore, it is possible that negative halo effects are indirectly contributing to the differential sex prevalence rates of ADHD and ODD.

There are several limitations in the current study. First, raters were undergraduate college students, and many of them may have had limited experience with child behavior. In addition, these students were 85% Caucasian, so the results cannot be generalized to other ethnic groups. Future studies should include parents or teachers as raters before the rater bias hypothesis can be ruled out as an explanation for differential sex prevalence rates in ADHD. In addition, a more ethnically diverse sample of raters would allow for greater generalization of the findings. In a similar vein, research is also needed in the area of ethnic differences in the assessment of children (Gingerich, Turnock, Litfin, & Rosén, 1998). In addition to including a more ethnically diverse sample of raters, future studies might include ethnically diverse children in the vignettes.

Although previous studies of negative halo effects have employed teachers as raters, sex differences in ratings of boys and girls displaying identical behaviors has not previously been evaluated. Thus, the current methodology may have produced different effects if teachers and/or parents had been the raters. Second, the vignettes used in the study tended to depict pure and relatively severe cases of each disorder. Using pure depictions of the disorders assists in isolating effects with statistical analyses.

However, pure cases of ADHD and/or ODD may be the exception rather than the rule. Therefore, these depictions may limit the ecological validity of the findings. In addition, the lack of inclusion of a vignette describing a typical child with no significant symptomatology could have led raters to form the impression that all vignettes would involve children with significant behavior and/or emotional problems. Therefore, future studies should include an asymptomatic child as well as a child with comorbid symptomatology. Although comorbidity presents a challenge for statistical analyses, it would provide a better representation of the types of clients who are seen in clinical practice.

In summary, the current study did not support the rater bias hypothesis as a partial explanation for differential sex prevalence rates in ADHD and ODD because no sex differences in the levels of symptomatology were found. However, future studies should employ a methodology to tests for sex differences in ratings of boys and girls who are displaying less severe levels of symptomatology. With regard to negative halo effects, the current study supports the existence of a bidirectional effect in boys and girls, as found by Jackson and King (2004) with teacher ratings, and extends the finding to ratings from noneducators. Although there were no sex differences in the magnitudes of these halo effects, they could indirectly contribute to the differential sex prevalence rate found in ADHD and ODD. Because ADHD and ODD are more common in boys than in girls, negative halo effects might have an exponential effect on prevalence estimates for boys as compared to girls. However, the clinical significance of these findings is still unclear because we do not know if the negative halo effect finding will generalize beyond written vignettes and videotaped scenarios to ratings of actual children. More research is needed in the area of sex biases in the assessment of children using rating scales so that the limitations of these measures, and the differential sex prevalence rate for ADHD and ODD, can be more fully understood. In addition, it is hoped that this study will lead to future research that will have implications for clinicians who assess children for disruptive behavior disorders.

References

Abikoff, H., Courtney, M., Pelham, W. E., & Koplewicz, H. S. (1993). Teachers' ratings of disruptive behaviors: The influence of halo effects. *Journal of Abnormal Child Psychology*, *21*, 519-533.

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

Barkley, R. A. (1998). *Attention-Deficit Hyperactivity Disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guildford.

Barkley, R. A. (2000). *Taking charge of ADHD: The complete, authoritative guide for parents* (rev. ed.). New York: Guildford.

Biederman, J., Mick, E., Faraone, S. V., Braaten, E., Doyle, A., Spencer, T., et al. (2002). Influence of gender on attention deficit hyperactivity disorder in children referred to a psychiatric clinic. *American Journal of Psychiatry*, *159*, 36-42.

DeFries, J. (1989). Gender ratios in children with reading disability and their affected relatives: A commentary. *Journal of Learning Disabilities*, *22*, 534-545.

Eme, R. (1992). Selective female affliction in the developmental disorders of childhood: A literature review. *Journal of Clinical Child Psychology*, *21*, 354-364.

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

Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders*, *5*, 143-154.

Gingerich, K. J., Turnock, P., Litfin, J. K., & Rosén, L. A. (1998). Diversity and attention deficit hyperactivity disorder. *Journal of Clinical Psychology*, *54*, 415-426.

Hartung, C. M., & Widiger, T. A. (1998). Sex differences in the diagnosis of mental disorders: Conclusions and controversies of the *DSM-IV. Psychological Bulletin*, *123*, 260-278.

Jackson, D. A., & King, A. R. (2004). Gender differences in the effects of oppositional behavior on teacher ratings of ADHD symptoms. *Journal of Abnormal Child Psychology*, *32*, 215-224.

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

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin*, *121*, 95-113.

Oltmanns, T. F., Neale, J. M., & Davidson, G. D. (1999). *Case studies in abnormal psychology* (5th ed.). New York: John Wiley.

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.

Pisecco, S., Huzinee, C., & Curtis, D. (2001). The effects of child characteristics on teachers' acceptability of classroom-based behavioral strategies and psychostimulant medication for the treatment of ADHD. *Journal of Clinical Child Psychology*, *30*, 413-421.

Rhee, S. H., Waldman, I. D., Hay, D. A., & Levy, F. (2001). Aetiology of the sex difference in the prevalence of *DSM-III-R* ADHD: A comparison of two models. In F. Levy & D. A. Hay (Eds.), *Attention, genes, and ADHD* (pp. 139-156). East Sussex, UK: Brunner-Routledge.

Schachar, R., Sandberg, S., & Rutter, M. (1986). Agreement between teachers' ratings and observations of hyperactivity, inattentiveness, and defiance. *Journal of Abnormal Child Psychology*, *14*, 331-345.

Sciutto, M. J., Nolfi, C. J., & Bluhm, C. (2004). Effects of child gender and symptom type on referrals for ADHD by elementary school teachers. *Journal of Emotional and Behavioral Disorders*, *12*, 247-253.

Stevens, J., Quittner, A. L., & Abikoff, H. (1998). Factors influencing elementary school teachers' ratings of ADHD and ODD behaviors. *Journal of Clinical Child Psychology*, *27*, 406-414.

Cynthia M. Hartung, PhD, graduated with a BS in psychology from the University of Wisconsin–Madison. She completed her doctoral studies in clinical psychology in the Department of Psychology at the University of Kentucky. In addition, she completed a clinical internship in the Department of Psychiatry at the University of Chicago Medical Center, a postdoctoral clinical fellowship at the Child Development and Rehabilitation Center at the Oregon Health Sciences University, and a postdoctoral research fellowship in the Departments of Psychology and Psychiatry at the University of Denver and the University of Colorado Health Sciences Center, respectively. Currently, she is an assistant professor in the Department of Psychology at Oklahoma State University. Her research interests include evidence-based assessment and treatment of disruptive behavior disorders, sex differences in developmental psychopathology, and prevention of peer relationship problems including bullying and victimization.

Jill C. Van Pelt, PhD, graduated with a BA in psychology and an MA in experimental psychology from Wake Forest University. She received her doctoral degree in clinical psychology in the Department of Psychology at Oklahoma State University in 2005. In addition, she completed a clinical internship at the University of Minnesota Medical School. Currently, she is a postdoctoral fellow in pediatric neuroscience at the University of Minnesota Medical School.

Monica L. Armendariz, MA, graduated with a BA in psychology from New Mexico State University. She received her master's degree in clinical psychology in the Department of Psychology at Oklahoma State University in 2005. Currently, she is a doctoral student in clinical psychology at Oklahoma State University.

Laura A. Knight, PhD, graduated with a BA in theater from Roger Williams College. In addition, she received an MA in psychology from Roosevelt University and an MS in clinical psychology from Oklahoma State University. She received her doctoral degree in clinical psychology in the Department of Psychology at Oklahoma State University in 2005. In addition, she completed a clinical internship at the University of Minnesota Medical School. Currently, she is a postdoctoral fellow in the Department of Child and Adolescent Psychiatry at the University of Michigan.