Published in final edited form as:
J Abnorm Child Psychol. 2009 July ; 37(5): 653–664. doi:10.1007/s10802-009-9309-1.

The Role of Harsh Discipline in Explaining Sex Differences in Conduct Disorder: A Study of Opposite-Sex Twin Pairs

Madeline H. Meier,
Department of Psychological Sciences and Midwest Alcoholism Research Center, University of Missouri-Columbia

Wendy S. Slutske,
Department of Psychological Sciences and Midwest Alcoholism Research Center, University of Missouri-Columbia

Andrew C. Heath, and
Washington University School of Medicine

Nicholas G. Martin
Queensland Institute of Medical Research

Abstract

In the current study, two hypotheses about the role of harsh discipline (HD) in explaining the sex difference in the prevalence of conduct disorder (CD) were evaluated: that boys exhibit more CD than girls because (1) they are exposed to more HD and/or (2) there is a greater association between HD and CD in boys. These hypotheses were evaluated in a sample of male and female adult twins from different families (N=3,502) as well as a sample of adult twin brothers and sisters (N=655) in order to examine the extent to which sex differences remained after controlling for between-family differences. Retrospective reports of HD experienced between ages 6–13 and DSM-IV CD symptoms experienced before age 18 were obtained via structured psychiatric telephone interviews. Boys reported higher mean levels of HD and CD than girls, both between and within families, and the results of model-fitting analyses suggested that differences in the use of harsh disciplinary practices for sons versus daughters may partially explain the sex difference in the prevalence of CD. Between families, the relation between HD and CD was greater for girls than boys, but within families, there was no evidence of a sex difference in the relation between HD and CD. Inconsistent between-family and within-family results suggest that factors that differ between families are confounded with sex differences in the relation between HD and CD. A more stringent test of sex differences involves eliminating these between-family differences by studying boys and girls within the same family.

Keywords

Conduct disorder; harsh discipline; sex differences

Males engage in more antisocial behavior than females, and this male preponderance appears to be ubiquitous. That is, regardless of developmental age, type of antisocial behavior (e.g., violent vs. non-violent behavior), or how antisocial behavior is assessed (self-report, parent-report, official crime records), it is predominantly males who engage in antisocial behavior, with few exceptions (Moffitt et al. 2001; Robins, 1991). Discovery of

Correspondence to: M. H. Meier. Department of Psychological Sciences, University of Missouri-Columbia, 210 McAlester Hall, Columbia, MO 65211, USA. mah4zc@mizzou.edu.

the causes of this sex disparity in the prevalence of antisocial behavior is critical to understanding the etiology of antisocial behavior (Keenan, Loeber, & Green, 1999; Lahey et al., 2006; Moffitt et al., 2001; Rutter, Caspi, & Moffitt, 2003), yet there are few empirical explanations for this sex difference.

Males and females share many of the same risk factors for antisocial behavior (Keenan et al., 1999; Moffitt et al., 2001; Rowe, Vazsonyi, & Flannery, 1995). However, males and females may be differentially exposed or differentially vulnerable to these risk factors, which may account for the male preponderance of antisocial behavior (Moffitt et al., 2001). That is, if males experience greater mean levels of a risk factor (greater exposure) and/or the association between a specific risk factor and antisocial behavior is greater for males than females (greater vulnerability) then this differential exposure or vulnerability could explain why antisocial behavior is more common among males and less common among females (Moffitt et al., 2001).

In exploring these two hypotheses – that males' differential exposure and/or vulnerability to risk for antisocial behavior may account for their higher rate of antisocial behavior – particular risk factors for antisocial behavior are arguably more likely to prove fruitful for study than others. For example, harsh discipline (HD; excessive corporal/physical punishment) is an excellent candidate risk factor for study of the mechanisms involved in the sex difference in the prevalence of antisocial behavior, because research suggests that males are more harshly disciplined than females (Lytton & Romney, 1991; Moffitt et al., 2001), and HD is associated with antisocial behavior within each sex (Deater-Deckard & Dodge, 1997; Gershoff, 2002; Moffitt et al., 2001; Rothbaum & Weisz, 1994).

There are many theories about how and why HD and antisocial behavior are related (Gershoff, 2002). A number of theories posit that HD may cause antisocial behavior. For example, harshly disciplined children may imitate parental aggression. Additionally, HD may inhibit the development of a close bond between parents and children, preventing children from internalizing their parents' values and conventional societal values and increasing their risk for deviant behavior (Gershoff, 2002). Alternatively, child antisocial behavior may educe HD. A number of genetically informative studies have found that the majority of the association between negative parenting and adolescent antisocial behavior can be explained by genetic factors, which is suggestive of an evocative gene-environment correlation (Jaffee, Caspi, Moffitt, Polo-Tomas, Price, & Taylor, 2004; Narusyte, Andershed, Neiderhiser, & Lichtenstein, 2007; Neiderhiser, Reiss, Hetherington, & Plomin, 1999; Pike, McGuire, Hetherington, Reiss, & Plomin, 1996). However, children may be more likely to elicit milder, normative forms of corporal punishment (e.g., spanking) than they are to provoke physical maltreatment (e.g., punching, kicking; Jaffee, Caspi, Moffitt, Polo-Tomas, Price, & Taylor, 2004; Jaffee, Caspi, Moffitt, & Taylor, 2004). Risk for physical maltreatment may depend more on characteristics of the parent and other factors that differ between families than on child behavior (Jaffee, Caspi, Moffitt, Polo-Tomas et al., 2004; Jaffee et al., 2004). Finally, the relation between parenting and child behavior is probably reciprocal, with child antisocial behavior eliciting physical discipline which in turn escalates the antisocial behavior (Pardini, Fite, & Burke, 2008; Patterson, Reid, & Dishion, 1992).

Several studies have examined sex differences in exposure and/or vulnerability to HD, but few of them have tested whether these sex differences can account, at least partially, for the male preponderance of antisocial behavior. Only two studies have addressed whether a sex difference in mean levels of HD can account for the sex difference in the prevalence of conduct problems. Moffitt et al. (2001) showed that greater levels of HD among boys than girls at ages 7 and 9 accounted for 6% of the mean-level sex difference in lifetime diagnoses

of conduct disorder (CD), and Messer, Goodman, Rowe, Meltzer, and Maughan (2006) found that 3.5% of the sex difference in disruptive behavior disorder diagnoses (CD, oppositional defiant disorder, or disruptive behavior disorder not otherwise specified) was associated with a greater likelihood of receiving physical punishment among boys than girls in a cross-sectional study of 5–10 year olds.

While boys may receive more HD than girls, it is not clear that there is a sex difference in the magnitude of the association between HD and antisocial behavior. Some studies have reported no sex difference (Dodge, Pettit, Bates, & Valente, 1995; Moffitt et al., 2001), others have reported that the relation between HD and antisocial behavior is greater for girls (Farrington & Painter, 2006; Messer et al., 2006), and still others that the relation is greater for boys (Gershoff, 2002; Leve, Kim, & Pears, 2005; Rothbaum & Weisz, 1994). The reason for these mixed findings is unclear, but they may be a result of the common practice of studying boys and girls from different families. There is some evidence that between-family differences can confound sex differences. For example, Kendler, Thornton, and Prescott (2001) found that the association between death of a friend and depression was greater among women than men in a sample of men and women from different families. However, within families, there was no sex difference in the association between death of a friend and depression. Such findings highlight the possibility that between-family differences are confounded with a sex difference in the relation between HD and antisocial behavior. Perhaps because of mixed findings regarding a sex difference in the relation between HD and antisocial behavior, no studies have tested whether sex differences in vulnerability to HD can account for the sex difference in the prevalence of antisocial behavior.

There were two aims of the present study. The first aim was to test whether and to what extent sex differences in exposure and vulnerability to HD can account for the male preponderance of CD. The second aim was to determine whether between-family differences confound sex differences in exposure and/or vulnerability to HD. These questions were examined in a large sample of male, female, and opposite-sex twin pairs. Sex differences in exposure and vulnerability to HD were examined between families using one twin randomly selected from each twin pair and within families using opposite-sex twin pairs. The use of opposite-sex twin pairs represents an optimal design for exploring sex differences (Moffitt et al., 2001; Rutter et al., 2003), because it involves the comparison of boys and girls that are matched on important characteristics, such as age, and family background factors, such as socioeconomic status, parental characteristics, and neighborhood of residence. Opposite-sex twin pairs are also partially matched on genetic factors. This design has the methodological advantage of eliminating between-family confounds that arise in studies of boys and girls from different families as well as offering greater statistical power to detect sex differences (Barnett, Marshall, Raudenbush, & Brennan, 1993; Rutter et al., 2003). To our knowledge, this represents the first investigation of the sex difference in antisocial behavior that has capitalized on the many advantages of studying opposite-sex twin pairs.

Method

Participants

Participants were members of the Australian Twin Registry, a community-based, volunteer twin registry. Details concerning participant recruitment are reported elsewhere (Lynskey et al., 2002). This data collection was approved by the Ethics Committee of the Queensland Institute of Medical Research and the IRB at the Washington University School of Medicine. The use of the data for this project was approved by the IRB at the University of Missouri-Columbia.

Data for this study were obtained from twins born between 1964–1971. The Semi-Structured Assessment for the Genetics of Alcoholism (Bucholz et al., 1994), which includes a diagnostic assessment of DSM-IV CD as well as an assessment of childhood family environment, was adapted for telephone use for the purposes of this study. These psychiatric interviews were conducted with the twins in 1996–2000 when the twins were between 24 and 36 years of age (M=30, SD=2.44). The participation rate for this interview was 84%. In all, twins from 3,502 families were sampled. For between-family analyses, one twin was randomly selected from each twin family. Fifty-three percent of the randomly selected participants for the between-family analyses were female. Due to missing data, 89–92% of twin families were useable for between-family analyses. For within-family analyses, only opposite-sex twin pairs for which both brother and sister had complete data were included in the analyses. Thus, 89–95% of the total sample of 655 opposite-sex twin pairs was used.

Demographic data for the opposite-sex twins are provided below. The majority (86%) of opposite-sex twins were of European ancestry, and the remaining opposite-sex twins were of aboriginal Australian ancestry (<1%), Asian ancestry (<1%) or mixed ancestry (13%). Most opposite-sex twins (69%) lived primarily in a large city in Australia during their childhood, whereas 19% lived in a small town and 12% lived in a rural area. Excluding their cotwin, twins from opposite-sex twin pairs had, on average, approximately two additional siblings (M=2.19, median = 2, range = 0–12). The opposite-sex twins retrospectively reported on the socioeconomic standing of their childhood families; 7% reported that their family was financially better off, 75% reported that their family was about average, and 18% reported that their family was financially worse off compared to other families in their childhood community.

Measures

Harsh discipline—Each twin retrospectively reported on how they were parented between ages 6 and 13, answering the following six items concerning the usual manner in which they were punished: 1) “Was the way your parents punished you mild or harsh?” (Figure 1: ‘harsh’), 2) “When you did something wrong were you often smacked by your parents?” (Figure 1: ‘smack’), 3) “Were your parents generally fair in scolding or punishing you or were they too lenient or too harsh?” (Figure 1: ‘too harsh’), 4) “When you did something wrong were you often punched or hit with a belt or stick by your parents?” (Figure 1: ‘hit’), 5) “How often did your parents physically punish you so hard that you hurt the next day?” (Figure 1: ‘hurt’), and 6) “What was the usual way in which your parents punished you?” (Figure 1: ‘usual way’). The response options for the first three HD items were dichotomous (i.e., mild/harsh, yes/no, fair or lenient/too harsh). Items four and five were measured on a 4-point scale (i.e., never, rarely, sometimes, often). Item six was measured on a 3-point scale (i.e., non-physical mild punishment [e.g., scold, isolate, fine], physical mild punishment [e.g., slap, spank], and physical harsh punishment [e.g., use weapon, punch, kick]). The items were scored such that higher values indicated greater levels of HD. We used these items to develop a dimensional, latent variable representing HD. In developing the measurement model of HD, the first three items were treated as categorical indicators of latent HD, and the last three items were treated as continuous indicators. We also fitted a model in which all items were treated as categorical or ordered categorical indicators, and we found that the model was not adversely affected by treating the last three items as continuous indicators. The coefficient alpha for the six HD items was 0.81, and the four-year test-retest reliability of a six-item HD scale among 215 participants in the current study who were randomly selected for a retest interview was r=0.73. Additionally, the intraclass correlation coefficient for HD among 2,037 same-sex twin pairs in the current study was 0.53.

Conduct Disorder—Participants retrospectively reported on the 15 DSM-IV (Diagnostic and Statistical Manual of Mental Disorders – Version IV) symptoms of CD they experienced before age 18. Each symptom of CD was assessed for both seriousness and pervasiveness. These symptoms were used to develop a dimensional, latent variable representing propensity to engage in conduct disordered behavior. This approach to the measurement of psychiatric disorders in general, and CD in particular, has received empirical support (Moffitt et al., 2008). Evidence suggests that the cut-off point of three symptoms for a diagnosis of CD is arbitrary, and variation above and below this cut-point provides important prognostic information (Moffitt et al., 2008). In developing our measurement model of CD, four symptoms were dropped due to low base rates: forced sexual activity, forced stealing, cruelty to animals, and staying out late at night before age 13. In addition, the symptom assessing fire setting with the intent to cause serious damage was also dropped due to low base rates. This symptom was replaced in our model by the following symptom: "Did you ever deliberately set fires when you were not supposed to?" This symptom loaded highly on the latent CD trait, indicating that this item correlates well with the other CD symptoms. The coefficient alpha for the 11 CD symptoms was 0.65, and the four-year test-retest reliability of an 11-item CD symptoms scale was r=0.76.

Data Analysis

Structural equation modeling was employed to test sex differences in exposure and vulnerability to HD. The use of structural equation modeling provided a more rigorous test of sex differences by incorporating an evaluation of the comparability of the measurement of HD and CD across boys and girls (i.e., measurement invariance). Tests of measurement invariance are particularly important in studies of sex differences because invariant measures rule out the possibility that any sex differences observed are due to differences in the measures, rather than in the constructs of interest. Structural equation modeling is also advantageous because it allows for the modeling of measurement error. Tests of measurement invariance were conducted first followed by tests of the exposure and vulnerability hypotheses.

Measurement invariance of each construct, HD and CD, was tested separately in a series of two steps (Byrne, Shavelson, & Muthén, 1989). In the first step, baseline models of the construct were estimated simultaneously for boys and girls, allowing measurement parameters (factor loadings, thresholds/intercepts) to be sex-specific. In the second step, the models of the construct were fit simultaneously for boys and girls and measurement parameters were constrained to be equal across sex. The fit of the constrained model was compared to the fit of the unconstrained model via a $\Delta\chi^2$ test. A non-significant $\Delta\chi^2$ test would suggest that the assumption of measurement invariance of the construct across sex was not rejected (i.e., the data were consistent with measurement invariance).

After establishing measurement invariance of the HD and CD constructs for boys and girls, tests of the exposure and vulnerability hypotheses were conducted. To test the exposure hypothesis, the mean-level sex difference in latent CD was estimated before and after controlling for latent HD. Specifically, two models were compared: 1) a model of the mean-level sex difference in latent CD based on the sex-invariant measurement model (the model of CD symptoms depicted in the bottom half of Figure 1) and 2) a model of the mean-level sex difference in latent CD in which HD predicted CD for boys and girls (Figure 1). In this second model, the parameter for the association between HD and CD was constrained to be equal across boys and girls (parameters 'A' and 'B' in Figure 1). Imposition of this constraint represented a presumption of no sex difference in vulnerability to HD, which would be tested later. This allowed us to isolate the effect of a sex difference in exposure to HD on CD, rather than estimating the combined effects of sex differences in exposure and

vulnerability to HD on CD. The extent to which the mean-level sex difference in CD diminished after controlling for HD would indicate the extent to which the sex difference in mean levels of CD could be accounted for by the sex difference in mean levels of HD.

To test the vulnerability hypothesis, two models were compared. In the first model, the association between HD and CD was freely estimated for boys and girls (paths 'A' and 'B' in Figure 1). In the second, nested model, the association between HD and CD was constrained to be equal for boys and girls (paths 'A' and 'B' in Figure 1). A sex difference in vulnerability to HD would be indicated by a significant decrease in model fit resulting from constraining the association between HD and CD to be equal across sex. The extent to which the mean-level sex difference in CD diminished after controlling for the sex difference in the relation between HD and CD would indicate the extent to which the sex difference in mean levels of CD could be accounted for by the sex difference in vulnerability to HD.

Structural equation models were fit in M*plus* (Muthén & Muthén, 1998–2004) using the mean- and variance-adjusted weighted least squares estimator (WLSMV). Between families, the tests of measurement invariance and substantive hypotheses were conducted within a multiple-group framework in which group was designated by participants' sex. This framework allows for the comparison of unrelated boys and girls. The within-family analyses were identical to the between-family analyses, with the exception that a multiple-group framework was not employed. A multiple-group framework would not have allowed for the matched comparison of brothers and sisters. The differences between the between-family and within-family models are depicted in Figure 1. In the within-family model, covariances between brothers' and sisters' latent constructs were modeled (paths 'C' and 'D' in Figure 1), and the dependency that exists among brothers' and sisters' observations was accounted for via the residual covariance among brothers' and sisters' CD. In the between-family model, covariances between boys' and girls' latent constructs were not modeled (paths 'C' and 'D' in Figure 1), because there would be no reason to expect that there would be any association between unrelated boys' and girls' HD, for example.

Results

In both the between-family and within-family samples, boys were more likely to report HD than girls (see Table 1) and more likely to engage in conduct disordered behavior than girls (see Table 2).

Further, more boys than girls in our sample met DSM-IV diagnostic criteria for CD. Ten percent of the sample (girls: 4%, boys: 18%) met lifetime criteria for CD (the presence of 3 or more symptoms of CD before age 18 but not necessarily clustering within a 12-month period). Seven percent of the sample (girls: 3%, boys: 13%) met the more restrictive criteria for CD in which symptoms were required to cluster within a 12-month period. Finally, 4% of the sample (girls: 2%, boys: 7%) met full diagnostic criteria for CD (3 or more symptoms clustering within a 12-month period plus impairment).

Measurement Model and Mean-Level Sex Difference for HD

A one-factor model of HD was estimated simultaneously for boys and girls and measurement parameters (factor loadings, intercepts for continuous indicators, and thresholds for categorical indicators) were allowed to be sex-specific. For both the boys' and girls' model, the residual variances of two factor indicators were allowed to covary due to their similarity in content ("Were your parents generally fair in scolding or punishing you or were they too lenient or too harsh?" and "Was the way your parents punished you mild or harsh?"). Next, this one-factor model of HD was estimated simultaneously for boys and

girls, and measurement parameters were constrained to be equal across sex (between families: RMSEA=0.06, CFI=0.98, TLI=0.98; within families: RMSEA=0.05, CFI=0.96, TLI=0.98). The fit of this constrained model was not significantly poorer than the fit of the unconstrained model (between families: $\Delta\chi^2$ (4, *N*=3,112)=25.63, p=0.06; within families: $\Delta\chi^2$ (4, *N*=585)=6.25, p=0.18), indicating that the data were consistent with measurement invariance of the HD construct across boys and girls.

This invariant measurement model was used to estimate the mean-level sex difference in HD. Between families, the mean-level sex difference in HD was 0.16 (*d*=0.43), with boys reporting more HD than girls. This mean difference was statistically significant, as holding the boys' and girls' means equal resulted in a significant decrease in model fit ($\Delta\chi^2$(1, *N*=3,112) =29.10, p<0.001). Within families, the mean-level sex difference in HD was 0.15 (*d*=0.25), with brothers reporting more HD than their twin sisters. This mean difference was also statistically significant ($\Delta\chi^2$(1, *N*=584)=13.48, p<0.001).

Measurement Model and Mean-Level Sex Difference for CD

A one-factor model of CD symptoms was fit for boys and girls simultaneously, and measurement parameters (factor loadings and thresholds) were freely estimated for each sex. In the within-family model, the residual variance of each factor indicator for twin brothers was allowed to covary with the associated factor indicator for their twin sisters. Next, measurement parameters were constrained to be equal across sex (between families: RMSEA=0.02, CFI=0.95, TLI=0.96; within families: RMSEA=0.03, CFI=0.91, TLI=0.93). The fit of this constrained model was not significantly poorer than the fit of the unconstrained model (between families: $\Delta\chi^2$ (7, *N*=3,185)=10.19, p=0.18; within families: $\Delta\chi^2$ (6, *N*=620)=12.12, p=0.06), indicating that the data were consistent with measurement invariance of the CD construct across sex.

The invariant measurement model was used to estimate the mean-level sex difference in CD. Between families, the mean difference in CD for boys and girls was 1.63 (*d*=5.6), which indicated that boys reported more CD symptoms than girls. This mean difference was statistically significant ($\Delta\chi^2$ (1, *N*=3,185)=30.14, p<0.001). Within families, the mean-level sex difference in CD was 1.54 (*d*=3.0), which indicated that brothers reported more CD symptoms than their twin sisters. This mean difference was statistically significant ($\Delta\chi^2$ (1, *N*=620)=11.77, p<0.001).

The Effect of the Sex Difference in Exposure to HD on the Sex Difference in Mean Levels of CD

After determining that boys were significantly more likely to report HD than girls and that boys also reported higher levels of CD than girls both between and within families, we tested the extent to which the mean-level sex difference in HD could account for the mean-level sex difference in CD. To do this, we compared the sex difference in the mean level of CD before controlling for HD to the sex difference in the mean level of CD obtained after controlling for HD. In this latter model, we assumed there was not an interaction between HD and sex in predicting CD - the paths from HD to CD were held equal for boys and girls (this assumption was tested later).

Between families, the sex difference in the mean level of CD before controlling for HD was 1.63. After controlling for HD, the sex difference in CD was 0.89. Comparison of these means revealed that the sex difference in the mean level of HD accounted for 45% ((1.63–0.89)/1.63) of the sex difference in the mean level of CD.

Within families, the sex difference in the mean level of CD before controlling for HD was 1.54. The sex difference in the mean level of CD after controlling for HD was 0.96.

Comparison of these means revealed that the sex difference in the mean level of HD accounted for 38% ((1.54–0.96)/1.54) of the sex difference in the mean level of CD.

The Effect of a Sex Difference in Vulnerability to HD on the Sex Difference in Mean Levels of CD

In the previous analyses, we assumed that the relation between HD and CD was the same for boys and girls. In the current analyses, we tested this assumption by comparing the previous model where the path from HD to CD was constrained to be equal across sex to a new model where this path was allowed to vary across sex.

Between families, the estimate for the association between HD and CD based on the constrained model was b=0.36 for boys and girls. When the association between HD and CD was allowed to vary across boys and girls from different families, the estimate was b=0.09 for boys and b=0.38 for girls. This unconstrained model fit the data well (RMSEA=0.03, CFI=0.97, TLI=0.98) and better than the constrained model ($\Delta\chi^2$ (1, N=3,103)=6.39, p=0.01), which indicated that the relation between HD and CD was greater for girls than boys. The sex difference in the mean level of CD, after taking into account both the higher mean level of HD among boys and the greater vulnerability to HD among girls, was 1.60. Comparison of this sex difference in the mean level of CD to the mean difference obtained before controlling for HD (1.63) indicated that sex differences in exposure and vulnerability to HD collectively accounted for 2% ((1.63–1.60)/1.63) of the between-family sex difference in the mean level of CD.

Within families, the constrained model equating the estimates for the association between HD and CD across brothers and sisters (RMSEA=0.02, CFI=0.96, TLI=0.97) did not fit the data significantly worse than the unconstrained model ($\Delta\chi^2$ (1, N=584)=1.76, p=0.19). This finding indicated that the relation between HD and CD was no different for brothers and sisters. Thus, sex differences in exposure but not vulnerability to HD could account for the sex difference in the mean level of CD. The results of the model constraining the association between HD and CD to be equal across brothers and sisters are presented in Figure 2.

The Effect of the Sex Difference in Mean Levels of CD on the Sex Difference in Mean Levels of HD

All of the previous structural models assumed a causal effect of HD on CD. We also tested the alternative structural model, whereby CD has a causal effect on HD. We present results based on the within-family sample only. Results revealed that the mean-level sex difference in HD (0.15) did not change after controlling for CD (0.15). Further, the causal effect of CD on HD (b=0.30) was no different for boys and girls ($\Delta\chi^2$ (1, N=584) =3.15, p=0.08). The fit of this model of the causal effect of CD on HD (RMSEA=0.02, CFI=0.97, TLI=0.97) was no different than the fit of the previous model of the causal effect of HD on CD (RMSEA=0.02, CFI=0.96, TLI=0.97).

Discussion

The first aim of this study was to test whether sex differences in exposure and/or vulnerability to HD could account for higher levels of CD among boys than girls, and the second aim of this study was to determine whether between-family differences were confounded with sex differences in exposure and vulnerability to risk for CD.

Consistent with previous research (Lytton & Romney, 1991, Moffitt et al., 2001), boys reported higher levels of HD and CD than girls. Evidence consistent with the exposure hypothesis was obtained in both the between- and within-family analyses. However, the between- and within-family approaches yielded different answers regarding the vulnerability

hypothesis. Between families, there was evidence that girls were more vulnerable to the effects of HD than boys. As boys were more harshly disciplined than girls but girls were more affected by HD than boys, sex differences in exposure and vulnerability to HD jointly accounted for only 2% of the mean-level sex difference in CD between families. Within families, sisters were not more vulnerable to the effects of HD than their brothers. Although brothers were more harshly disciplined than sisters, there was no evidence of a sex difference in vulnerability to HD. Thus, higher levels of HD among brothers than sisters alone accounted for 38% of the mean-level sex difference in CD within families.

The discrepant between-family and within-family results suggest that sex differences in vulnerability to HD may be confounded with between-family differences. Boys and girls from different families differ on many factors, such as family socioeconomic status, parental personality, and number of additional siblings. Differences in some of these factors across boys and girls from different families may explain why girls appeared to be more vulnerable to HD between but not within families. Within families, boys and girls are matched on these factors, allowing for a more sensitive, rigorous test of sex differences. Results of this study suggest that the inconsistent pattern of results found in the literature regarding sex differences in the relation between HD and antisocial behavior may be explainable by between-family differences that are confounded with sex differences.

An important strength of this study is that it does not suffer from a bias that characterizes most of the research on sex differences in HD and CD, that is, failing to account for between family differences that might contribute to sex differences. However, there are also several limitations of this study. Reports of HD and CD were obtained retrospectively and may be subject to reporting bias. For example, Henry, Moffitt, Caspi, Langley, and Silva (1994) found evidence of both underreporting and overreporting of delinquent acts. Though prospective reports are certainly preferred, retrospective reports should not be dismissed altogether. Evidence suggests that adults still can provide important information about childhood experiences, especially when the reports are used to test hypotheses regarding relative standing rather than precise event dates and frequencies (Henry et al., 1994) and when those experiences are relatively concrete and objective (Brewin, Andrews, & Gotlib, 1993; Hardt & Rutter, 2004). Though our measure of HD included some clearly ‘subjective’ items, the four-year test-retest reliability of this measure was adequate. Further, the intraclass correlation coefficient among same-sex twins for HD was impressively large, given that one would expect that parents treat their children somewhat differently. Another limitation is that the retrospective reports of HD and CD were obtained using the same informant, which may lead to inflated estimates of their association.

It could be argued that sex differences in retrospective recall coupled with shared method variance may at least partially explain the results of this study. This explanation, although plausible, seems somewhat unlikely in light of the findings. In terms of the exposure hypothesis, we found that boys reported slightly more harsh discipline than girls, and this finding is consistent with past research (Lytton & Romney, 1991; Moffitt et al., 2001). With respect to the vulnerability hypothesis, we found no evidence of a sex difference in the relation between HD and CD. If a sex difference in retrospective recall exists, it would have to be in the opposite direction of the true effect to produce this null finding. For example, if the association between HD and CD were greater among girls than boys, retrospective reports among males would have to result in an inflated estimate of the association between HD and CD in order to produce a finding of no sex difference in this association.

An additional limitation of this study is the assumption that parental HD causes CD. An alternative explanation is that conduct disordered children elicit HD or that CD and HD reciprocally influence each other. The nature of the sex difference in the relation between

HD and CD may depend on the direction of causation modeled. For example, it may be that boys are more likely to elicit HD, potentially because they exhibit more conduct disorder symptoms, but girls are more vulnerable to the effects of HD. If this were the case, failure to model the bidirectional influences would result in an overall finding of no sex difference in the relation between HD and CD.

We attempted to distinguish between the various directional explanations via direction-of-causation modeling (Gillespie, Zhu, Neale, Heath, & Martin, 2003; Heath et al., 1993). Structural equation models fitted to twin data can provide information about the direction of causation of the association between two traits under certain restrictive circumstances, for example, when one trait is largely influenced by genetic factors and the other is largely influenced by family environmental factors. Unfortunately, these restrictive circumstances were not met for HD and CD, and the models testing these various hypotheses were indistinguishable based on model fit.

The directional relation between HD and CD remains an unresolved question in the extant literature, though it seems likely that the relation is bidirectional (Lochman, 2008). Our decision to model the causal effect of HD on CD was influenced by evidence suggesting that harsh physical discipline, compared to corporal punishment, may be less dependent on child behavior and more determined by parental characteristics (Jaffee, Caspi, Moffitt, & Taylor, 2004; Jaffee, Caspi, Moffitt, Polo-Tomas et al., 2004). That is, child misbehavior may elicit punishment, such as a spanking, from parents, but it is factors associated with the parent or the environment that determine whether the child will be punished with a punch or kick. Thus, results of our study must be understood in light of the assumption that HD causes CD.

Additionally, it must be noted that there may be something unique about sex differences within opposite-sex twin pairs. Sex differences in opposite-sex twin pairs may be muted if, for example, girls from opposite-sex twin pairs were influenced by intrauterine male hormones. Resnick, Gottesman, and McGue (1993) compared opposite-sex to same-sex twins and found that girls from opposite-sex twin pairs had higher levels of sensation seeking than girls from same-sex twin pairs. This finding suggests that girls from opposite-sex twin pairs may be masculinized due to prenatal exposure to male androgens, though psychosocial explanations are possible as well. For example, imitation effects have been reported for antisocial behavior (Carey, 1992), and girls from opposite-sex twin pairs may imitate their male co-twin. Imitation can occur in all families however, and it could be that the presence of a male sibling (whether or not a twin or a sibling close in age) is associated with increased antisocial behavior among girls. Thus, it is not clear if there is something unique about sex differences within opposite-sex twin pairs. A comparison of opposite-sex twins and opposite-sex, non-twin siblings could clarify this issue.

Finally, the twins in our sample grew up in Australia, are of mostly European ancestry, and came from average socioeconomic backgrounds, and it is possible that our findings will not generalize across all cultures or sociodemographic groups. However, our finding of harsher parental disciplinary practices for boys than girls is consistent with findings reported by Lytton & Romney (1991) who found that physical punishment was administered more to boys than girls, and this sex difference was most pronounced in Western countries. Further, the prevalence of CD for boys and girls in our sample was consistent with reports of the prevalence of CD in community-based samples from different countries (Loeber, Burke, Lahey, Winters, & Zera, 2000).

In summary, the results of this study suggest that the greater prevalence of CD among boys compared to girls arises, in part, because boys are exposed to more HD and not because boys are more vulnerable to HD than girls. These conclusions are consistent with those

reported by Moffitt et al. (2001). However, whereas Moffitt et al. (2001) reported that the sex difference in exposure to HD accounted for 6% of the sex difference in CD, our findings suggest a much greater role (38%) for exposure to HD in explaining the sex difference in CD. The reason for this difference in effect size is difficult to isolate, as our study differed greatly from this prior study in many ways, including sample, measures of HD and CD, and data analytic approach. Nonetheless, the results suggest that disciplining boys more similarly to girls may reduce CD symptomatology in boys by as much as 38%.

The study of opposite-sex twin pairs is a promising approach to discovering the causes of the sex disparity in antisocial behavior, though a sample of opposite-sex, non-twin siblings would also provide many of the same advantages as opposite-sex twin pairs. Directions for future research will be to use a prospective, multi-informant research design that can potentially disentangle the causal relation between HD and CD and to examine sex differences in the effects of other putative risk factors on CD.

Acknowledgments

Data collection for this study was funded by National Institutes of Health grant AA010249. Preparation of this article was supported in part by National Institutes of Mental Health Grant F31MH079683. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health.

Abbreviations

HD	harsh discipline
CD	conduct disorder

References

Barnett RC, Marshall NL, Raudenbush SW, Brennan RT. Gender and the relationship between job experiences and psychological distress: A study of dual-earner couples. Journal of Personality and Social Psychology. 1993; 64:794–806. [PubMed: 8505708]

Brewin CR, Andrews A, Gotlib IH. Psychopathology and early experience: A reappraisal of retrospective reports. Psychological Bulletin. 1993; 113:82–98. [PubMed: 8426875]

Bucholz KK, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI, Reich T, Schmidt I, Schuckit MA. A new, semi-structured psychiatric interview for use in genetic linkage studies: a report of the reliability of the SSAGA. Journal of Studies on Alcohol. 1994; 55:149–158. [PubMed: 8189735]

Byrne BM, Shavelson RJ, Muthén B. Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin. 1989; 105:456–466.

Carey G. Twin imitation for antisocial behavior: Implications for genetic and family environment research. Journal of Abnormal Psychology. 1992; 101:18–25. [PubMed: 1537964]

Deater-Deckard K, Dodge KA. Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. Psychological Inquiry. 1997; 8:161–175.

Dodge KA, Pettit GS, Bates JE, Valente E. Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. Journal of Abnormal Child Psychology. 1995; 104:632–643.

Farrington, DP.; Painter, KA. Gender differences in offending: Implications for risk-focused prevention. 2004. Retrieved March 1, 2006 from http://www.homeoffice.gov.uk/rds/pdfs2/rdsolr0904.pdf

Gershoff ET. Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. Psychological Bulletin. 2002; 128:539–579. [PubMed: 12081081]

Gillespie NA, Zhu G, Neale MC, Heath AC, Martin NG. Direction of causation modeling between cross-sectional measures of parenting and psychological distress in female twins. Behavior Genetics. 2003; 33:383–396. [PubMed: 14574138]

Hardt J, Rutter M. Validity of adult retrospective reports of adverse child experiences: A review of the evidence. Journal of Child Psychology and Psychiatry. 2004; 45:260–273. [PubMed: 14982240]

Heath AC, Kessler RC, Neale MC, Hewitt JK, Eaves LJ, Kendler KS. Testing hypotheses about direction of causation using cross-sectional family data. Behavior Genetics. 1993; 23:29–50. [PubMed: 8476389]

Henry B, Moffitt TE, Caspi A, Langley J, Silva PA. On the “remembrance of things past”: A longitudinal evaluation of the retrospective method. Psychological Assessment. 1994; 2:92–101.

Jaffee SR, Caspi A, Moffitt TE, Taylor A. Physical maltreatment victim to antisocial child: Evidence of an environmentally mediated process. Journal of Abnormal Psychology. 2004; 113:44–55. [PubMed: 14992656]

Jaffee SR, Caspi A, Moffitt TE, Polo-Tomas M, Price TS, Taylor A. The limits of child effects: Evidence for genetically mediated child effects on corporal punishment but not on physical maltreatment. Developmental Psychology. 2004; 6:1047–1058. [PubMed: 15535755]

Keenan K, Loeber R, Green S. Conduct disorder in girls: A review of the literature. Clinical Child and Family Psychology Review. 1999; 2:3–19. [PubMed: 11324095]

Kendler KS, Thornton LM, Prescott CA. Gender differences in the rates of exposure to stressful life events and sensitivity to their depressogenic effects. American Journal of Psychiatry. 2001; 158:587–593. [PubMed: 11282693]

Lahey BB, Van Hulle CA, Waldman ID, Rodgers JL, D’Onofrio BM, Pedlow S, Rathouz P, Keenan K. Testing descriptive hypotheses regarding sex differences in the development of conduct problems and delinquency. Journal of Abnormal Child Psychology. 2006; 34:737–755. [PubMed: 17033935]

Leve LD, Kim HK, Pears KC. Childhood temperament and family environment as predictors of internalizing and externalizing trajectories from ages 5 to 17. Journal of Abnormal Child Psychology. 2005; 33:505–520. [PubMed: 16195947]

Lochman JE. Bidirectional parent-child relationships. [Special issue]. Journal of Abnormal Child Psychology. 2008; 36(5)

Loeber R, Burke JD, Lahey BB, Winters A, Zera M. Oppositional and conduct disorder: A review of the past 10 years, part I. Journal of the American Academy of Child and Adolescent Psychiatry. 2000; 39:1468–1484. [PubMed: 11128323]

Lynskey MT, Heath AC, Nelson EC, Bucholz KK, Madden PAF, Slutske WS, Statham DJ, Martin NG. Genetic and environmental contributions to cannabis dependence in a national young adult twin sample. Psychological Medicine. 2002; 32:195–207. [PubMed: 11871371]

Lytton H, Romney DM. Parents’ differential socialization of boys and girls: A meta-analysis. Psychological Bulletin. 1991; 109:267–296.

Messer J, Goodman R, Rowe R, Meltzer H, Maughan B. Preadolescent conduct problems in girls and boys. Journal of the American Academy of Child and Adolescent Psychiatry. 2006; 45:184–191. [PubMed: 16429089]

Moffitt TE, Arseneault L, Jaffee SR, Kim-Cohen J, Koenen KC, Odgers CL, Slutske WS, Viding E. Research review: DSM-V conduct disorder: research needs for an evidence base. Journal of Child Psychology and Psychiatry. 2008; 49:3–33. [PubMed: 18181878]

Moffitt, TE.; Caspi, A.; Rutter, M.; Silva, PA. Sex differences in antisocial behavior: Conduct disorder, delinquency, and violence in the Dunedin Longitudinal Study. New York, NY: Cambridge University Press; 2001.

Muthén, LK.; Muthén, B. Mplus User’s Guide. Los Angeles, CA: Author; 1998–2004.

Narusyte J, Andershed AK, Neiderhiser JM, Lichtenstein P. Aggression as a mediator of genetic contributions to the association between negative parent-child relationships and adolescent antisocial behavior. European Child and Adolescent Psychiatry. 2007; 16:128–37. [PubMed: 17136502]

Neiderhiser JM, Reiss D, Hetherington EM, Plomin R. Relationships between parenting and adolescent adjustment over time: genetic and environmental contributions. Developmental Psychology. 1999; 35:680–692. [PubMed: 10380859]

Pardini DA, Fite PJ, Burke JD. Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: the moderating effect of age and African-American ethnicity. Journal of Abnormal Child Psychology. 2008; 36:647–662. [PubMed: 17899362]

Patterson, GR.; Reid, JB.; Dishion, TJ. Antisocial boys. Vol. IV. Eugene, OR: Castalia; 1992. A social learning approach.

Pike A, McGuire S, Hetherington E, Reiss D, Plomin R. Family environment and adolescent depressive symptoms and antisocial behavior: a multivariate genetic analysis. Developmental Psychology. 1996; 32:590–604.

Resnick SM, Irving I, McGue M. Sensation seeking in opposite-sex twins: An effect of prenatal hormones? Behavior Genetics. 1993; 23:323–329. [PubMed: 8240211]

Robins LN. Conduct disorder. Journal of Child Psychology and Psychiatry. 1991; 32:193–209. [PubMed: 2037645]

Rothbaum F, Weisz JR. Parental caregiving and child externalizing behavior in nonclinical samples: A meta-analysis. Psychological Bulletin. 1994; 116:55–74. [PubMed: 8078975]

Rowe DC, Vazsonyi AT, Flannery DJ. Sex differences in crime: Do means and within-sex variation have similar causes? Journal of Research in Crime & Delinquency. 1995; 32:84–100.

Rutter M, Caspi A, Moffitt TE. Using sex differences in psychopathology to study causal mechanisms: Unifying issues and research strategies. Journal of Child Psychology and Psychiatry. 2003; 44:1092–1115. [PubMed: 14626453]

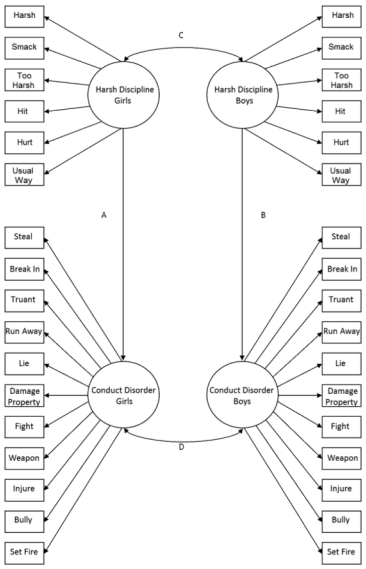

Figure 1.

Structural equation model testing sex differences in exposure to HD and sex differences in the relation between HD and CD.

Circles represent the unobserved, latent constructs of CD and HD. Squares represent the observed behaviors that are influenced by and indicators of the latent constructs. Single-headed arrows represent causal paths, and double-headed arrows represent covariances. The top half of the figure depicts the measurement model for HD. The bottom half of the figure depicts the measurement model for CD. Paths “A” and “B” represent the association between HD and CD for girls and boys, respectively. Paths “C” and “D” represent the covariance between girls’ and boys’ HD and CD constructs within families, respectively. The model illustrated is for within-family analyses. The between-family model is identical, except that paths “C” and “D” are not included. For simplicity, the mean structure is not depicted.

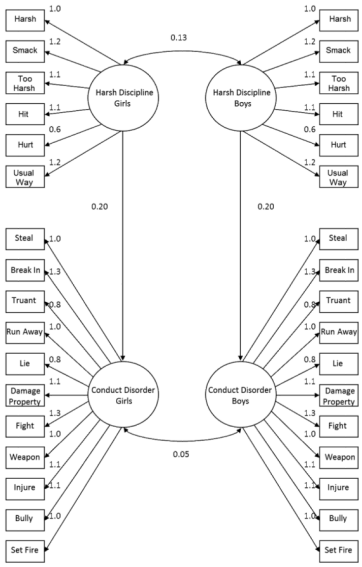

Figure 2.

Structural equation model testing sex differences in exposure to HD and sex differences in the relation between HD and CD within families. The numbers associated with the paths leading from the latent variables to the observed variables represent factor loadings. The path coefficients associated with the causal paths leading from HD to CD indicate the amount of change in CD associated with a unit increase in HD. All parameter estimates are unstandardized. For simplicity, the mean structure and residual covariances are not presented. This model fit the data well (RMSEA=0.02, CFI=0.96, TLI=0.97).

Table 1

Sex differences in HD between and within families

	Between Families		Within Families	
Item	**Boys (%)**	**Girls (%)**	**Brothers (%)**	**Sisters (%)**
Was the way your parents punished or disciplined you mild or harsh? (Harsh)	21	22	20	22
When you did something wrong, were you often smacked by your parents? (Yes)	51	45*	52	43*
Were your parents generally fair in scolding or punishing you or were they too lenient or too harsh? (Too harsh)	12	15*	12	15
When you did something wrong, were you ever punched or hit with a belt or stick or something like that by your parents? (Yes)	67	48*	65	47*
Did you parents ever physically punish you so hard that you hurt the next day? (Yes)	16	15	19	13*
What was the usual way in which your parents punished or disciplined you? (Used weapon, punch, kick)	30	18*	28	20*

Note: Between families: Boys N = 1,388, Girls N = 1,724. Within families: Boys N = 584, Girls N = 584.

* $p<0.01$

J Abnorm Child Psychol. Author manuscript; available in PMC 2011 April 26.

Table 2

Sex differences in CD symptoms between and within families

	Between Families		Within Families	
Symptom	**Boys (%)**	**Girls (%)**	**Brothers (%)**	**Sisters (%)**
Truant	6	3*	6	2*
Run away	3	3	3	4
Lie/con	17	10*	18	12*
Steal	4	3	4	3
Damage property	15	2*	17	3*
Physical fights	12	2*	12	2*
Use a weapon	4	2*	4	2
Injure someone	10	3*	10	4*
Bully	3	2	3	2
Light fire	25	6*	25	6*
Break into car/house	12	2*	13	1*

Note: Between families: Boys N = 1,423, Girls N = 1,762. Within families: Boys N = 584, Girls N = 584.

* p<0.01