HHS Public Access

Author manuscript

Psychol Med. Author manuscript; available in PMC 2020 May 18.

Published in final edited form as:

Psychol Med. 2019 July ; 49(10): 1600–1607. doi:10.1017/S0033291719000771.

How Should We Understand the Absence of Sex Differences in the Genetic and Environmental Origins of Antisocial Behavior?

S. Alexandra Burt, Ph.D.[1], **Brooke L. Slawinski, M.A.**[1], **E. Elisa Carsten, B.A.**[1,2], **K. Paige Harden, Ph.D.**[3], **Luke W. Hyde, Ph.D.**[4], **Kelly L. Klump, Ph.D.**[1]

[1]Department of Psychology, Michigan State University

[2]Department of Psychology, University of South Florida

[3]Department of Psychology, University of Texas at Austin

[4]Department of Psychology, University of Michigan

Abstract

Available twin-family data on sex differences in antisocial behavior (ASB) simultaneously suggest that ASB is far more prevalent in males than in females, and that its etiology (i.e., the effects of genes, environments, hormones, culture) does not differ across sex. This duality presents a conundrum: How do we make sense of mean sex differences in ASB if not via differences in genes, environments, hormones, and/or cultures? The current selective review and critique explores possible contributions to these seemingly incompatible sets of findings. We asked whether the presence of sex differences in behavior could be smaller than is typically assumed, or confined to a specific set of behaviors. We also asked whether there might be undetected differences in etiology across sex in twin-family studies. We found little evidence that bias or measurement invariance across sex account for phenotypic sex differences in ASB, but we did identify some key limitations to current twin-family approaches. These included the questionable ability of qualitative sex difference analyses to detect gender norms and prenatal exposure to testosterone, and concerns regarding specific analytic components of quantitative sex difference analyses. We conclude that the male preponderance in ASB is likely to reflect a true sex difference in observed behavior. It was less clear, however, that the genetic and environmental contributions to ASB are indeed identical across sex, as argued by prior twin-family studies. It is our hope that this review will inspire the development of new, genetically-informed methods for studying sex differences in etiology.

Antisocial behavior (ASB) is defined as actions that violate societal norms and the personal or property rights of others, and includes both overt or aggressive behaviors (fighting, hitting, bullying) and covert or non-aggressive/rule-breaking behaviors (stealing, lying, vandalism). Consistent with modern operationalizations of psychopathology, ASB is

Correspondence: Address correspondence to Alex Burt, Department of Psychology, Michigan State University, 107D Psychology Building, East Lansing, MI 48824. Electronic mail may be sent to burts@msu.edu. Fax number is (517) 432-2476.

Conflict of Interest: None

Ethical Standards: The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

operationalized here as a continuous trait. Males engage in significantly more ASB than do females beginning in the toddler years and continuing throughout the lifespan, with typical Cohen's *d* effect sizes ranging from 0.4 to 0.8 (Archer, 2004, Hyde, 1984, Knight *et al.*, 1996). Mean sex[1] differences are especially large for severe violence, with observed odds ratios of 18.8 and higher (e.g., Daly and Wilson, 1990, van Lier *et al.*, 2009). They are smaller, but still moderate-to-large in magnitude, for more garden-variety acts of physical aggression (Archer, 2004, Hyde, 1984, Knight *et al.*, 1996), and small-to-moderate in magnitude for non-aggressive ASB (Burt, 2012, Moffitt, 2003). This male preponderance persists across numerous human societies (Archer, 2004, Ramirez *et al.*, 2001) and across most mammalian species, including humans' nearest phylogenetic cousin, the chimpanzee (Gray, 1971, Maccoby and Jacklin, 1980, Manson *et al.*, 1991).

Given these seemingly clear sex differences in ASB at the level of actual behavior, one might expect equally robust differences in genetic and environmental etiology of ASB across sex. One possible manifestation involves 'qualitative sex differences' in etiology, in which different genes and/or different environments influence ASB in males versus females. As detailed in Table 1, qualitative sex differences in etiology should theoretically act to decrease the similarity of opposite-sex siblings relative to same-sex siblings, since (for example) hormonal effects that contribute to ASB only in boys would degrade similarity between male and female siblings and/or increase same-sex sibling similarity. Consistent with this possibility, at least one twin study (Rose *et al.*, 2004) found evidence of sex-specific genetic influences on ASB. These results are bolstered by a recent GWAS (Tielbeek *et al.*, 2017), although these results were based on a relatively small sample by GWAS-standards (N=8,535 females and 7,772 males). That said, these findings appear to be exceptions to the rule. The majority of twin studies, and all adoption studies, have found no/minimal evidence of qualitative sex differences in ASB etiology (e.g., Eaves *et al.*, 1997, Jacobson *et al.*, 2002, van Hulle *et al.*, 2007), an important set of null findings given that twin-family designs are able to simultaneously evaluate both genetic and environmental influences (something that cannot be done using molecular genetic designs). Jacobson and colleagues (2002), for example, evaluated this possibility in a large longitudinal study with over 1,000 opposite sex twin pairs, comparing opposite-sex twin covariances to same-sex twin covariances in qualitative sex limitation models. They found no evidence that the same-sex dizygotic twin correlation statistically exceeded that for opposite-sex dizygotic twins in their models, arguing against qualitative sex differences in the etiology of ASB. Such results are compelling, given the large sample size of opposite-sex pairs, but we note that no meta-analysis examining the question of qualitative sex differences in ASB has ever been conducted. While such work is needed before any firm conclusions can be drawn, extant data generally suggest that genetic and environmental contributions to ASB do not differ across males and females.

Another potential manifestation of sex differences in etiology is quantitative in nature. Findings of quantitative sex differences would indicate that, although the specific genes and

[1]We are focused here on biological sex, as determined by the X and Y chromosomes, rather than gender identity, which is a separate but important construct that requires additional research. Though many of the studies cited herein studies are likely measuring "gender" via self-report, our discussion is framed in terms of biological sex.

environments influencing ASB do not differ across sex (i.e., there are no qualitative differences), the magnitude(s) of those influences do differ, such that they are more important for one sex versus the other (see Table 1). Unfortunately, results regarding quantitative sex differences in ASB etiology are quite inconsistent across the literature. Some studies have reported clear evidence against quantitative sex differences in etiology (e.g., Burt *et al.*, 2001, Eaves *et al.*, 1997, Gelhorn *et al.*, 2005, Slutske *et al.*, 1997, Taylor *et al.*, 2000), while others have reported clear evidence for quantitative sex differences (e.g., Bartels *et al.*, 2003, Eley *et al.*, 1999, Jacobson *et al.*, 2002, Rose *et al.*, 2004, Silberg *et al.*, 1994, Van den Oord *et al.*, 1994). Critically, however, some of the latter found ASB to be more heritable in males (Bartels *et al.*, 2003, Silberg *et al.*, 1994, Van den Oord *et al.*, 1994), while others found that ASB was more heritable in females (e.g., Eley *et al.*, 1999, Jacobson *et al.*, 2002, Rose *et al.*, 2004).

Given these inconsistencies, it is perhaps not surprising that the question of quantitative sex differences in the etiology of ASB has also been addressed as part of several large-scale meta-analyses of twin and adoption studies. These studies have uniformly suggested that ASB is equally heritable in males and females (Burt, 2009a, b, Rhee and Waldman, 2002). When the original meta-analytic data analyzed in Burt (2009a) were additionally disambiguated by *informant* (Burt *et al.*, 2018), however, the conclusion changed: rather than no quantitative sex differences in etiology at all, the data indicated that there were no sex differences in etiology when ASB was assessed using maternal informant-reports of child behavior. When examining teacher informant-reports of child ASB, however, ASB was more shared environmental in origin in boys than in girls, but more genetic in origin in girls than in boys – a conclusion that persisted to an independent twin sample not included in the original meta-analysis (Burt *et al.*, 2018). Such findings are thought to reflect the 'attribution bias context model', whereby mothers and teachers are exposed to different slices of the child's behavior and thus develop different opinions/attributions regarding the same child (De Los Reyes and Kazdin, 2005). For example, ASB might be more gendered in its presentation at school (perhaps due to child-enforced social norms that constrain physical aggression with peers in girls) and less gendered in the home, where mothers may observe their daughter hitting her sibling(s). Alternately, it may be the case that ASB on the part of any one child is simply easier to observe in scholastic settings because it is so disruptive to the functioning of the classroom, or because the teacher has a large pool of children to which they can compare any given child. Either way, the etiological differences observed for teacher informant-reports, but not maternal informant-reports, suggest that quantitative sex differences in the etiology of ASB are specific to (or more detectable in) school settings, a peculiar finding given that mean differences in ASB are seen across all informants.

What do the above findings mean for our understanding of sex differences in ASB?

Although interesting, this specificity of etiologic differences to teacher-informant reports in particular still leaves us with something of a conundrum. When examining maternal informant-reports of their children (by far the most frequently examined of the informant-reports), available data robustly suggest that ASB is far more prevalent in boys as compared

to girls, while also indicating that neither the magnitude nor the composition of its etiology vary across boys and girls. In more specific terms, extant data indicate that *1)* the same genetic and environmental influences underlie maternal informant-reports of ASB in males and females, *2)* these genetic and environmental influences are equally influential in males and females, and yet *3)* ASB is notably more common in males than in females (a phenotypic difference observed for both maternal and teacher-informant-reports). How do we make sense of these seemingly incompatible sets of findings?

Before we dive in to possible answers, it is worth highlighting one key distinction between phenotypic and etiologic sex differences. Discussions of phenotypic sex differences are almost exclusively based on mean differences across sex, such that average levels of ASB are higher in males than they are in females. This does not imply, however, that all or even most males evidence higher rates of ASB than do females, since there is a great deal of variability around each of these sex-specific means[2]. Discussions of etiology, by contrast, are based on decompositions of the variance (or observed individual differences) around the mean, without regard to the mean itself. Put another way, phenotypic and twin studies are respectively focused on different statistical moments (i.e., means versus variances). As such, the fact that their conclusions are difficult to reconcile is frustrating but does not necessarily imply that either is incorrect (although they could be, and indeed, there are nuanced issues for both that warrant additional consideration, as discussed below).

Despite this statistical truism, our core question remains unanswered: How do we make sense of mean sex differences in ASB if not via differences in genes, environments, hormones, and/or culture? The current review will consider a variety of answers to this question, ultimately highlighting areas where new theoretical and empirical designs could yield novel and much needed insights.

1. Could phenotypic sex differences in ASB be less pronounced than we think?

One possible contribution to the interpretive mismatch between sex differences in phenotype and sex differences in etiology could be that phenotypic differences are in fact smaller than is typically reported. Indeed, this concern has been raised, both in more general terms and for ASB in particular. In terms of the former, a large body of work has strongly suggested that, while there may be moderately-sized mean sex differences in ASB, most psychological constructs show only minimal sex differences, if any (see review in Hyde, 2005). For ASB more specifically, it has previously been argued that sex differences in physical aggression were due a failure to assess aggressive behaviors that are more salient for females, and particularly those that used interpersonal relationships to harm others (e.g., Crick and Grotpeter, 1995). Namely, it could be the case that girls' heavy focus on interpersonal relationships and social functioning means that relational aggression is their preferred weapon. Girls' smaller physical size and reduced physical strength may accentuate this tendency by limiting their capacity to use physical aggression effectively (Björkqvist, 1994).

[2]Boys and men also typically evidence larger ASB variances than girls/women. The presence of larger observed variances in males is still compatible with the absence of sex differences in biometric decompositions of that variance, however, since genetic and environmental parameter estimates are typically standardized (i.e., converted into proportions) for each sex prior to conducting comparisons across sex. See Table 1 for more details.

More recent meta-analytic work, however, has indicated that this early (and very influential) 'mean girl' ASB hypothesis may not be correct, since boys engage in just as much relational aggression as do girls (Card *et al.*, 2008). In short, there is as yet scant evidence that phenotypic sex differences in ASB stem from sex biases inherent to its operationalization.

It would also be important to meaningfully consider the issue of measurement invariance across sex, or the assumption that the underlying structure and properties of a measure are consistent across males and females. Measurement invariance evaluates whether the same construct is being measured across groups, and can be tested several ways. Configural invariance models, for example, would evaluate whether the factor structure of ASB is equivalent across sex, while metric invariance models would evaluate whether the ASB factor loadings are equivalent across sex. Evidence against measurement equivalence across sex would thus indicate that any observed sex differences were actually artifacts of measurement (e.g., informants using the scale differently for males and females) rather than true mean differences across sex. As it happens, however, empirical research robustly supports measurement invariance across sex for measures of ASB, and does so across developmental periods, informants, community/clinical settings, and nationalities (e.g., Fonseca-Pedrero *et al.*, 2012, Palmieri and Smith, 2007, Smits *et al.*, 2016). Such findings argue against the possibility that observed sex differences in ASB stem from measurement invariance across sex.

Although the sexually dimorphic nature of ASB may not be attributable to bias or measurement invariance, it could conceivably reflect other aspects of its operationalization, including the omission of underlying emotional and cognitive processes. ASB arises from a complex interaction of deficits across inhibitory control, reward and punishment sensitivity, emotionality and emotion regulation, and empathy (Hyde *et al.*, 2013, Waller *et al.*, 2015). The failure to observe sex differences in these underlying traits and cognitive processes could thus suggest that sex differences in ASB are less pronounced than they appear (although this is not the only possible interpretation). To date, however, there is at best inconsistent evidence to support this conjecture. Some (but not all) of these more basic psychological processes do appear to show sex differences, though the evidence for these differences is sometimes mixed and they are often small-to-moderate in magnitude. For example, boys are considerably less empathic than girls when empathy is measured by questionnaire, although these sex difference shrink when empathy is measured by physiology or observation (Eisenberg and Lennon, 1983). Impulsivity, a well-documented predictor of both the onset and development of ASB, demonstrates robust sex differences, such that males are more impulsive and risk-taking than females as well as less responsive to punishment across the lifespan (Cross *et al.*, 2011). That said, delay discounting, a narrower measure of reward sensitivity, demonstrates either no sex difference (Cross *et al.*, 2011) or one of very small effect ($d = .06$) (Silverman, 2003). Similarly, high Negative Emotionality/Neuroticism is also a known predictor of ASB, but is more pronounced in females than in males (Schmitt *et al.*, 2008), although this female preponderance is far more pronounced for anxiety and other internalizing traits than for those related to ASB (i.e., hostility). In short, the sexually dimorphic nature of ASB extends to several but not all of its core emotional and cognitive predictors.

Diving deeper, we could also ask whether there are consistent sex differences in the neural circuits underlying ASB (Alegria *et al.*, 2016). Unfortunately, much of this research has focused on male participants in prison settings (e.g., Crooks *et al.*, 2018b), and the handful of studies that have included females (e.g., Crooks *et al.*, 2018a, Fairchild *et al.*, 2013) have not included males, making it difficult to draw substantive conclusions. That said, results from the few community studies containing men and women thus far suggest little (or insufficient) evidence for sex differences in the neural correlates of ASB (Carré *et al.*, 2012, e.g., Dotterer *et al.*, 2017, though see Waller *et al.*, 2016). Interestingly, however, when we instead focus on the basic functioning of specific regions of interest to ASB (e.g., regions involved in emotion and reward), consistent sex differences do emerge. Men have smaller amygdala volumes (Lenroot and Giedd, 2010, Ruigrok *et al.*, 2014), reduced amygdalar responsivity to negative stimuli (Stevens and Hamann, 2012), and different patterns of neural reactivity to reward (Dreher *et al.*, 2007). As an example of this duality, Hyde and colleagues (2014) found that sex predicted threat-related neural activity in the amygdala, while also finding no evidence that sex moderated associations between neural reactivity in those regions and ASB (Hyde *et al.*, 2014). In sum, extant data indicate that, although there may be sex differences in brain volume and activation in the structure and function of neural regions related to ASB, it is not the case that (for example) amygdalar activation is associated with ASB more in one sex versus the other. That said, more work is needed before any firm conclusions can be drawn. Future studies should clarify whether (as we would assume) the aforementioned sex differences in brain volume and activation explain the preponderance of ASB in men as compared to women. There is also a need for studies using larger samples that directly examine whether sex moderates associations between neural reactivity and ASB. Pending those studies, however, we preliminarily conclude that, as with the personality and cognitive correlates discussed above, there is little evidence that the neural correlates of ASB differ across sex.

2. Are twin studies well-suited for the detection of qualitative sex differences?

Another possible contribution to the interpretive mismatch between phenotype and etiology could be that twin family studies may not be appropriate vehicles for evaluations of qualitative sex differences. One key hurdle in this regard relates to sex-specific cultural norms. Norms against aggression, for example, are especially salient for girls, who are socialized against this behavior by both adults and peers to a greater extent than boys (Keenan & Shaw, 1997; Crick et al., 2007). Critically, however, the effects of these (and other) sex-specific cultural norms are detectable in twin studies only under particular scenarios. One such scenario would occur if individuals within a given sex are differentially responsive to cultural norms based on their genetic predilections to ASB (e.g., if women with a higher genetic loading for ASB are more or less responsive to cultural gender norms than women without that loading). Because responsiveness to the cultural norm depends on one's genetic influences in this case, the effect of sex-specific norms would actually be subsumed in the genetic component of variance in ASB, and should thus be detectable in qualitative sex difference analyses. Similarly, should individuals be differentially exposed to ASB-relevant cultural norms, it would also be detectable (likely within the shared environmental component of variance, since it involves exposure rather than genetically-influenced responsiveness).

The qualitative sex differences model would be quite useful for detecting sex differences in etiology under either those scenarios (both of which are quite reasonable a priori). Even so, there are other reasonable scenarios that would be more difficult to detect within a qualitative sex differences twin design. A key proviso of twin models is that they are blind to any factors that are shared by all persons in a sample; if something does not differ across a specific population, it cannot alter individual differences within that group. Put another way, it does not alter their rank ordering within the group (i.e., those highest in a given trait remain highest, etc). To the extent that norms against aggression (for example) are communicated more or less universally to any girl growing up in a particular place at a particular time, it is theoretically possible that they could suppress mean levels of ASB across all girls without affecting the rank-order of ASB among girls. Because the rank-ordering of girls remains unaffected under this scenario, this manifestation of cultural effects should not alter either same-sex or opposite-sex twin correlations. As such, we would be unable to detect the norm's effect in qualitative sex difference analyses despite its clear sex-specific effect on ASB.

Yet another possible hurdle for the interpretation of qualitative sex difference analyses relates to gonadal hormone exposure (which is ironic given that sex-specific gonadal hormones often motivate studies of qualitative sex differences). Gonadal hormones are foundational to the sexual differentiation of humans. It is the presence of prenatal testosterone in males, and the relative absence of prenatal testosterone in females, that lead to male and female physiology, respectively. These organizational effects of prenatal testosterone also permanently alter brain structure and function, and in doing so, contribute to many different types of sex differentiated behaviors, including aggression (Ryan and Vandenbergh, 2002, Spencer *et al.*, 2017). When considered alongside the presence of strong and very early sex differences in aggression (the peak prevalence of aggression is between the ages of 2 and 4 years), it thus seems likely that prenatal testosterone may contribute both to the etiology of ASB and to the male predominance in risk. There is one key caveat to these findings, however: within normal limits of prenatal testosterone exposure, there are at most modest associations between level of exposure and behavioral outcomes in males (Ryan and Vandenbergh, 2002, Tapp *et al.*, 2011). Despite prominent sex-specific effects on ASB, it is thus quite plausible that prenatal testosterone exposure elevates risk for ASB in all males without altering the rank ordering of males within groups (much like the final cultural norms possibility detailed above). Put another way, prenatal testosterone exposure could increase the mean level of ASB among males without differentially affecting opposite-sex sibling similarity and same-sex sibling similarity. If so, prenatal testosterone effects could be quite difficult to detect in typical qualitative twin studies.

What's more, some data suggest that the female co-twin in opposite-sex pairs may be exposed to higher levels of testosterone *in utero* by her male co-twin, and is therefore more "masculinized" in her brain and behavior (Cohen-Bendahan *et al.*, 2004, Tapp *et al.*, 2011) and aggression (Cohen-Bendahan *et al.*, 2005) than girls from same-sex pairs. Should this be true, it would also act increase the similarity between male and female co-twins from opposite sex pairs to levels more similar to those of same-sex pairs, thereby attenuating estimates of qualitative sex differences. In sum, under a few particular scenarios, it is

possible that null findings in qualitative sex difference analyses (i.e., the absence of significant sex differences) may represent false negatives.

3. Are twin studies well-suited for the detection of quantitative sex differences?

Less optimal elements inherent to current quantitative sex difference approaches may also contribute to the interpretive mismatch between phenotype and etiology. As outlined in Table 1, studies of quantitative sex differences in etiology are typically conducted as follows. First, we compare MZ and same-sex DZ covariances, separately for each sex, to determine standardized sex-specific heritability estimates. We then compare these sex-specific heritability estimates across sex, constraining parameter estimates to be equal across sex and evaluating changes in model fit. These constraint analyses require ample statistical power. As such, only large differences are likely to consistently emerge as statistically significant, an important consideration given the fact that even phenotypic sex differences are only moderate in magnitude in all but the most violent forms of ASB. Accordingly, at the level of the individual study, limited statistical power could account for the absence of consistent sex differences in etiology. This explanation falls short, however, when we consider the absence of sex differences in etiology in most meta-analytic studies, which typically rely on raw intraclass correlations (rather than summary statistics) to compute meta-analytic effect sizes.

The more important statistical concern with twin studies of quantitative sex differences may instead be the comparison of second-order genetic and environmental parameter estimates. Current statistical modelling approaches specify that genetic and environmental variances cannot be lower than zero (i.e., cannot be negatively-signed), thereby avoiding negative variances (as they are not possible in theoretical terms). This specification is necessary because genetic and environmental variance estimates are based on the comparison of MZ and DZ covariances, and can thus yield negative variances under a variety of circumstances (e.g., when the MZ correlation is more than double the DZ correlation, the estimated shared environmental variance would be negatively signed). As noted in recent work by Verhulst & Neale, however, this approach has some significant drawbacks, not least of which is that it leads to questionable Type I Error rates and introduces bias into all estimated parameters (since they are proportional to one another). They have thus developed a new approach that corrects the above issues by allowing for negative variances (Verhulst and Neale, 2018). As this new approach has yet to be applied to examinations of sex differences in the etiology, however, it remains unclear whether the absence of sex differences in the etiology of ASB will persist once unbounded analyses are conducted.

A related approach, and one that completely avoids the computation of higher-order genetic and environmental parameter estimates (but, to our knowledge, is rarely conducted), would be to formally compare intraclass correlations for a given zygosity across sex (e.g., compare male-male DZ similarity to female-female DZ twin similarity). This approach would have the significant advantage of increased statistical power relative to the comparisons of proportions of variance across sex, and could also enhance understanding. For example, should we find that MZ females are usually more similar than are MZ males, it would support early suggestions that girls require a high genetic liability to engage in ASB (Cloninger *et al.*, 1978). Similar approaches are now being applied to genotype-environment

interaction models, to promising effect (Turkheimer *et al.*, 2017). In sum, analyses that focus on intraclass correlations or allow for negative heritability estimates would increase our certainty that the magnitudes of genetic and environmental influences on ASB are equivalent across sex.

CONCLUSIONS

The above review was motivated by the following question: *How do we make sense of mean sex differences in ASB in the absence of sex differences in its etiology?* Our specific goal was to critically examine the overall question (e.g., could there be undetected differences in etiology across sex?), and in doing so, illuminate possible answers. We specifically explored the presence of true sex differences in ASB at the level of behavior, as well as the suitability of twin-family designs for qualitative and quantitative sex difference analyses, with the goal of gaming out reasonable scenarios under which the twin design would be unable to detect actual sex differences in the etiology of ASB.

Our review found little evidence that bias or measurement invariance across sex accounted for phenotypic sex differences. Similarly, several (but not all) relevant personological, cognitive, and emotional correlates appear to vary across sex in ways consistent with ASB. Based on this review, we thus conclude that the male preponderance of ASB is likely to reflect a true sex difference in observed behavior. By contrast, our exploration of the suitability of twin-family designs for uncovering sex differences in etiology yielded some intriguing possibilities. In terms of qualitative analyses, there are two potentially critical variables – sex-specific cultural norms and prenatal testosterone in males – that could conceivably alter mean levels of ASB (perhaps substantially so) without altering individual differences within a given sex. Under this particular scenario, sex-specific etiologic influences on ASB would be very difficult to detect using standard qualitative sex difference analyses (or quantitative analyses, for that matter). We also highlighted potential problems with the current approach to uncovering quantitative sex differences, most notably the central focus on the comparison of standardized heritability estimates. In short, although twin studies are widely used to uncover the presence of sex differences at the etiologic level, there are specific scenarios and specific types of sex differences that these models are unlikely to detect. The possibility of false negatives should give researchers some considerable pause when interpreting the results of null (or no-sex-difference) findings. Positive findings, on the other hand, seem more likely to reflect actual sex differences in etiology (be they quantitative or qualitative).

Given these results, efforts should now be made to explore and address potential limitations to twin study methods and interpretation. Our review indicated that current issues with quantitative sex difference analyses could likely be resolved with one or more relatively simple analytic tweaks (e.g., focusing on intraclass correlations, allowing variances to be negatively-signed). The concerns raised regarding qualitative sex difference analyses are more challenging to address, and as such, notably undermine our confidence in null results obtained using that model. Indeed, these concerns appear to be significant enough that they may require a considerable revamping of current qualitative approaches and/or the use of additional analyses to illuminate the interpretation of null results. As one example, we could

leverage the fact that different cultures have different cultural norms regarding female ASB to try and tease out sex-specific effects of culture on etiology, asking whether brother-sister similarity for ASB varies across communities or societies according to their gender norms (e.g., are opposite-sex twins/siblings more similar to one another in societies with less pronounced gendered norms or higher gender equality?). Migration studies (Breslau *et al.*, 2011) provide yet another strategy for examining cultural effects in general, since immigrants alter their environmental conditions without altering their genotypes. As an example of the overall logic, Breslau and colleagues (2011) examined Mexicans in various stages of migration to the United States: 1) Mexicans living in non-migrant households in Mexico, 2) Mexicans living in the United States as adults but who were raised in Mexico, as well as Mexicans living in Mexico but with an immediate family member living in the United States, 3) those of Mexican ancestry who were born in the United States or Mexicans who came to the United States as children, and 4) Mexican-Americans born in the United States to at least one US-born parent. Comparing migrants (i.e., group 2) to those born and/or raised in the new country (i.e., groups 3 and 4) allowed them to examine the influence of societal environmental conditions on behavior prevalence. Had they also examined these effects separately across sex, we could begin to evaluate gendered cultural effects on behavior prevalence as well. By contrast, to test theories regarding effects of prenatal testosterone, we may need to leverage naturally-occurring hormonal perturbations (e.g., Androgen Insensitivity Syndrome) and/or reinvigorate connections between human and animal behavioral genetic research (given that the latter allow for hormonal manipulations *in utero*). We are hopeful that the field will tackle these interesting possibilities.

As we close, we would like to reiterate a point made earlier on: the fact that the conclusions of phenotypic and etiologic studies are difficult to reconcile is frustrating but does not necessarily imply that either is incorrect. Even so, this review did undermine our confidence in null findings obtained using standard qualitative and quantitative biometric sex difference analyses. We hope the field will begin development of new methods for studying sex differences in etiology.

Financial Support:

This project was supported by an administrative supplement for R01-MH081813 from the National Institute of Mental Health (NIMH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIMH or the National Institutes of Health.

References

Alegria AA, Radua J & Rubia K (2016). Meta-analysis of fMRI studies of disruptive behavior disorders. American Journal of Psychiatry 173, 1119–1130. [PubMed: 27523497]

Archer J (2004). Sex differences in aggression in real-world settings: a meta-analytic review. Review of general Psychology 8, 291.

Bartels M, Hudziak J, Van den Oord E, Van Beijsterveldt C, Rietveld M & Boomsma D (2003). Co-occurrence of aggressive behavior and rule-breaking behavior at age 12: Multi-rater analyses. Behavior Genetics 33, 607–621. [PubMed: 14574136]

Björkqvist K (1994). Sex differences in physical, verbal, and indirect aggression: A review of recent research. Sex roles 30, 177–188.

Breslau J, Borges G, Saito N, Tancredi DJ, Benjet C, Hinton L, Kendler KS, Kravitz R, Vega W, Aguilar-Gaxiola S & Medina-Mora ME (2011). Migration From Mexico to the United States and

Conduct Disorder: A cross-national study Archives of General Psychiatry 68, 1284–1293. [PubMed: 22147845]

Burt SA (2009a). Are there meaningful etiological differences within antisocial behavior? Results of a meta-analysis. Clinical Psychology Review 29, 163–178. [PubMed: 19193479]

Burt SA (2009b). Rethinking environmental contributions to child and adolescent psychopathology: A meta-analysis of shared environmental influences. Psychological Bulletin 135, 608–637. [PubMed: 19586164]

Burt SA (2012). How do we optimally conceptualize the heterogeneity within antisocial behavior? An argument for aggressive versus non-aggressive behavioral dimensions. Clinical Psychology Review 32, 263–279. [PubMed: 22459789]

Burt SA, Krueger RF, McGue M & Iacono WG (2001). Sources of covariation among ADHD, CD, and ODD: The importance of shared environment. Journal of Abnormal Psychology 110, 516–525. [PubMed: 11727941]

Burt SA, Slawinski BL & Klump KL (2018). Are there sex differences in the etiology of youth antisocial behavior? Journal of abnormal psychology 127, 66. [PubMed: 29369668]

Card NA, Stucky BD, Sawalani GM & Little TD (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. Child Development 79, 1185–1229. [PubMed: 18826521]

Carré JM, Hyde LW, Neumann CS, Viding E & Hariri AR (2012). The neural signature of distinct psychopathic traits. Social Neuroscience 8, 122–135. [PubMed: 22775289]

Cloninger CR, Christiansen KO, Reich T & Gottesman II (1978). Implications of sex differences in the prevalences of antisocial personality, alcoholism, and criminality for familial transmission. Archives of General Psychiatry 35, 941–951. [PubMed: 354554]

Cohen-Bendahan CC, Buitelaar JK, van Goozen SH & Cohen-Kettenis PT (2004). Prenatal exposure to testosterone and functional cerebral lateralization: a study in same-sex and opposite-sex twin girls. Psychoneuroendocrinology 29, 911–916. [PubMed: 15177706]

Cohen-Bendahan CC, Buitelaar JK, Van Goozen SH, Orlebeke JF & Cohen-Kettenis PT (2005). Is there an effect of prenatal testosterone on aggression and other behavioral traits? A study comparing same-sex and opposite-sex twin girls. Hormones and Behavior 47, 230–237. [PubMed: 15664027]

Crick NR & Grotpeter JK (1995). Relational aggression, gender, and social-psychological adjustment. Child development, 710–722. [PubMed: 7789197]

Crooks D, Anderson NE, Widdows M, Petseva N, Decety J, Pluto C & Kiehl KA (2018a). The relationship between cavum septum pellucidum and psychopathic traits in female offenders. Behavioural Brain Research 112, 95–104.

Crooks D, Anderson NE, Widdows M, Petseva N, Koenigs M, Pluto C & Kiehl KA (2018b). The relationship between cavum septum pellucidum and psychopathic traits in a large forensic sample. Neuropsychologia 112, 95–104. [PubMed: 29545126]

Cross CP, Copping LT & Campbell A (2011). Sex differences in impulsivity: a meta-analysis. Psychological bulletin 137, 97. [PubMed: 21219058]

Daly M & Wilson M (1990). Killing the competition. Human Nature 1, 81–107. [PubMed: 24222024]

De Los Reyes A & Kazdin AE (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. Psychological Bulletin 131, 483–509. [PubMed: 16060799]

Dotterer HL, Hyde LW, Swartz JR, Hariri AR & Williamson DE (2017). Amygdala reactivity predicts adolescent antisocial behavior but not callous-unemotional traits. Developmental Cognitive Neuroscience 24, 84–92. [PubMed: 28279916]

Dreher J-C, Schmidt PJ, Kohn P, Furman D, Rubinow D & Berman KF (2007). Menstrual cycle phase modulates reward-related neural function in women. Proceedings of the National Academy of Sciences 104, 2465–2470.

Eaves LJ, Silberg JL, J.M. M, Maes HH, Simonoff E, Pickles A, Rutter M, Neale M, Reynolds CA, Erikson MT, Heath AC, Loeber R, Truett KR & Hewitt J (1997). Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the

Virginia twin study of adolescent development. Journal of Child Psychology and Psychiatry 38, 965–980. [PubMed: 9413795]

Eisenberg N & Lennon R (1983). Sex differences in empathy and related capacities. Psychological bulletin 94, 100–131.

Eley TC, Lichtenstein P & Stevenson J (1999). Sex differences in the etiology of aggressive and nonaggressive antisocial behavior: Results from two twin studies. Child Development 70, 155–168. [PubMed: 10191520]

Fairchild G, Hagan CC, Walsh ND, Passamonti L, Calder AJ & Goodyer IM (2013). Brain structure abnormalities in adolescent girls with conduct disorder. Journal of Child Psychology and Psychiatry 54, 86–95. [PubMed: 23082797]

Fonseca-Pedrero E, Sierra-Baigrie S, Lemos-GirÂldez S, Paino M & Muñiz J (2012). Dimensional structure and measurement invariance of the Youth Self-Report across gender and age. Journal of Adolescent Health 50, 148–153. [PubMed: 22265110]

Gelhorn HL, Stallings MC, Young SE, Corley RP, Rhee SH & Hewitt JK (2005). Genetic and environmental influences on conduct disorder: Symptom, domain, and full-scale analyses. Journal of Child Psychology and Psychiatry 46, 580–591. [PubMed: 15877764]

Gray JA (1971). Sex differences in emotional behaviour in mammals including man: endocrine bases. Acta psychologica 35, 29–46. [PubMed: 4995340]

Hyde JS (1984). How large are gender differences in aggression? A developmental meta-analysis. Developmental psychology 20, 722.

Hyde JS (2005). The gender similarities hypothesis. American psychologist 60, 581. [PubMed: 16173891]

Hyde LW, Byrd AL, Votruba-Drzal E, Hariri AR & Manuck SB (2014). Antisocial behavior and amygdala reactivity: Divergent correlates of antisocial personality and psychopathy traits in a community sample. Journal of Abnormal Psychology 123, 214–224. [PubMed: 24661171]

Hyde LW, Shaw DS & Hariri AR (2013). Neuroscience, developmental psychopathology and youth antisocial behavior: Review, integration, and directions for research. Developmental Review 33, 168–223.

Jacobson KC, Prescott CA & Kendler KS (2002). Sex differences in the genetic and environmental influences on the development of antisocial behavior. Development and Psychopathology 14, 395–416. [PubMed: 12030698]

Knight GP, Fabes RA & Higgins DA (1996). Concerns about drawing causal inferences from meta-analyses: An example in the study of gender differences in aggression. Psychological bulletin 119, 410. [PubMed: 8668746]

Lenroot RK & Giedd JN (2010). Sex differences in the adolescent brain. Brain and cognition 72, 46–55. [PubMed: 19913969]

Maccoby EE & Jacklin CN (1980). Sex differences in aggression: A rejoinder and reprise. Child development, 964–980. [PubMed: 7471931]

Manson JH, Wrangham RW, Boone JL, Chapais B, Dunbar R, Ember CR, Irons W, Marchant L, McGrew W & Nishida T (1991). Intergroup Aggression in Chimpanzees and Humans [and Comments and Replies]. Current anthropology 32, 369–390.

Moffitt TE (2003). Life-course persistent and adolescence-limited antisocial behavior: A research review and a research agenda In The causes of conduct disorder and serious juvenile delinquency (ed. Lahey B, Moffitt TE and Caspi A), pp. 49–75. Guilford: New York.

Neale MC, Røysamb E & Jacobson K (2006). Multivariate genetic analysis of sex limitation and G× E interaction. Twin Research and Human Genetics 9, 481–489. [PubMed: 16899154]

Palmieri PA & Smith GC (2007). Examining the structural validity of the Strengths and Difficulties Questionnaire (SDQ) in a US sample of custodial grandmothers. Psychological assessment 19, 189. [PubMed: 17563200]

Ramirez JM, Andreu JM & Fujihara T (2001). Cultural and sex differences in aggression: A comparison between Japanese and Spanish students using two different inventories. Aggressive behavior 27, 313–322.

Rhee S & Waldman ID (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. Psychological Bulletin 128, 490–529. [PubMed: 12002699]

Rose RJ, Dick DM, Viken RJ, Pulkkinen L & Kaprio J (2004). Genetic and environmental effects on conduct disorder and alcohol dependence symptoms and their covariation at age 14. Alcoholism: Clinical and Experimental Research 28, 1541–1548.

Ruigrok AN, Salimi-Khorshidi G, Lai M-C, Baron-Cohen S, Lombardo MV, Tait RJ & Suckling J (2014). A meta-analysis of sex differences in human brain structure. Neuroscience & Biobehavioral Reviews 39, 34–50. [PubMed: 24374381]

Ryan BC & Vandenbergh JG (2002). Intrauterine position effects. Neuroscience & Biobehavioral Reviews 26, 665–678. [PubMed: 12479841]

Schmitt DP, Realo A, Voracek M & Allik J (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. Journal of personality and social psychology 94, 168. [PubMed: 18179326]

Silberg JL, Erickson MT, Meyer JM, Eaves LJ, Rutter ML & Hewitt JK (1994). The application of structural equation modeling to maternal ratings of twins' behavioral and emotional problems. Journal of Consulting and Clinical Psychology 62, 510. [PubMed: 8063977]

Silverman IW (2003). Gender Differences in Delay of Gratification: A Meta-Analysis. Sex Roles 9, 451–463.

Slutske WS, Heath AC, Dinwiddie SH, Madden PAF, Bucholz KK, Dunne MP, Statham DJ & Martin NG (1997). Modeling genetic and environmental influences in the etiology of conduct disorder: A study of 2,682 adult twin pairs. Journal of Abnormal Psychology 106, 266–279. [PubMed: 9131847]

Smits IA, Theunissen MH, Reijneveld SA, Nauta MH & Timmerman ME (2016). Measurement Invariance of the Parent Version of the Strengths and Difficulties Questionnaire (SDQ) Across Community and Clinical Populations.

Spencer D, Pasterski V, Neufeld S, Glover V, O'connor TG, Hindmarsh PC, Hughes IA, Acerini CL & Hines M (2017). Prenatal androgen exposure and children's aggressive behavior and activity level. Hormones and behavior 96, 156–165. [PubMed: 28939371]

Stevens JS & Hamann S (2012). Sex differences in brain activation to emotional stimuli: a meta-analysis of neuroimaging studies. Neuropsychologia 50, 1578–1593. [PubMed: 22450197]

Tapp AL, Maybery MT & Whitehouse AJ (2011). Evaluating the twin testosterone transfer hypothesis: a review of the empirical evidence. Hormones and behavior 60, 713–722. [PubMed: 21893061]

Taylor J, McGue M & Iacono WG (2000). Sex differences, assortative mating, and cultural transmission effects on adolescent delinquency: A twin family study. Journal of Child Psychology and Psychiatry 41, 433–440. [PubMed: 10836673]

Tielbeek JJ, Johansson A, Polderman TJ, Rautiainen M-R, Jansen P, Taylor M, Tong X, Lu Q, Burt AS & Tiemeier H (2017). Genome-wide association studies of a broad spectrum of antisocial behavior. JAMA psychiatry 74, 1242–1250. [PubMed: 28979981]

Turkheimer E, Beam CR, Sundet JM & Tambs K (2017). Interaction between parental education and twin correlations for cognitive ability in a Norwegian conscript sample. Behavior genetics 47, 507–515. [PubMed: 28744605]

Van den Oord EJ, Boomsma DI & Verhulst FC (1994). A study of problem behaviors in 10-to 15-year-old biologically related and unrelated international adoptees. Behavior Genetics 24, 193–205. [PubMed: 7945150]

van Hulle CA, Rodgers JL, D'Onofrio BM, Waldman ID & Lahey BB (2007). Sex differences in the causes of self-reported adolescent delinquency. Journal of Abnormal Psychology 116, 236–248. [PubMed: 17516757]

van Lier PAC, Vitaro F, Barker ED, Koot HM & Tremblay RE (2009). Developmental links between trajectories of physical violence, vandalism, theft, and alcohol-drug use from childhood to adolescence. Journal of Abnormal Child Psychology 37, 481–492. [PubMed: 19089610]

Verhulst B & Neale M (2018). Type I error rates in multivariate behavioral genetic models. In Annual Meeting of the Behavior Genetics Association: Cambridge, MA.

Waller R, Corral-Frías NS, Vannucci B, Bogdan R, Knodt AR, Hariri AR & Hyde LW (2016). An oxytocin receptor polymorphism predicts amygdala reactivity and antisocial behavior in men. Social cognitive and affective neuroscience, nsw042.

Waller R, Murray L, Dotterer H & Hyde L (2015). Neuroimaging approaches to understanding youth antisocial behavior In Brain Mapping: An Encyclopedic Referenece (ed. Toga AW), pp. 1001–1009. Academic Press: Elsevier.

Table 1.

Type	Definition	Example	Required calculations
Qualitative Sex Differences	Different ACE factors influence ASB in males versus females (different genes and/or different environments, although the model cannot clarify which)	A given etiologic factor increases the likelihood of ASB only in one sex but not the other. This could include sex-specific genes like the Y chromosome, particular autosomal genes that matter for one sex but not the other (as in Tielbeek *et al.*, 2017), or gender-specific norms	• Compare DZ-OS to DZ-SS • Should covariances for DZ-SS be larger than those for DZ-OS, it implies the presence of sex-specific etiologic influences rDZ-SS > rDZ-OS • Note that, because there are no opposite-sex MZ twins, we can confirm the presence of qualitative sex differences in etiology, but not their genetic or environmental origin (Neale *et al.*, 2006)
Quantitative Sex Differences	The magnitude(s) of A, C, and/or E differ, such that they are more important in one sex versus the other	Although the genes and environments influencing ASB do not vary across sex, they account for more variance (i.e., matter more) in one sex versus the other. In scholastic contexts, for example, ASB is more genetic in origin in girls than in boys (Burt *et al.*, 2018).	• Compare MZ and DZ-SS covariances within each sex to compute unstandardized ACE estimates, separately by sex A = 2(rMZ-rDZ) C = (2*rDZ)-rMZ E = 1-rMZ • Standardize ACE estimates, within each sex A+C+E = 1.0 • Fit formal constraint models, constraining standardized ACE estimates for males and females, respectively, to be equal across sex, and evaluating reductions in model fit

Note. MZ, DZ-SS, and DZ-OS represent monozygotic, same-sex dizygotic, and opposite-sex dizygotic twins, respectively. rMZ and rDZ represent intraclass correlations for monozygotic and dizygotic twins, respectively. A, C, and E represent, additive genetic, shared environmental, and non-shared environmental variances, respectively.

Psychol Med. Author manuscript; available in PMC 2020 May 18.