SEX DIFFERENCES IN SOCIALLY AGGRESSIVE BEHAVIORS

By

Brooke L. Slawinski

A THESIS

Submitted to
Michigan State Univeristy
in partial fulfillment of the requirements
for the degree of

Psychology – Master of Arts

2016

ABSTRACT

SEX DIFFERENCES IN SOCIALLY AGGRESSIVE BEHAVIORS

By

Brooke L. Slawinski

Social aggression is often assumed to be the "girl form" of aggression, but previous investigations of sex differences have yielded notably inconsistent results. Unfortunately, there is no "gold standard" assessment for social aggression. Instead, there are many formal and informal measures, each of which contains unique behavioral items. This creates a lack of definitional and measurement clarity that may contribute to the inconsistent sex differences observed in prior work. It also raises an important question: is it possible that the inconsistencies in sex differences reported at the level of the overall scale are related to the specific items? Put another way, are there consistent sex differences for specific behaviors? Similarly, are previously reported sex differences due to true differences in the latent trait of social aggression or are they consequences of measurement bias? The current study addressed these possibilities by examining the presence and direction of sex differences in the perpetration of specific socially aggressive behaviors and testing whether observed sex differences are due to measurement non-invariance. Social aggression in middle childhood (N=2,000) and young adulthood (N=1,200) was assessed using the Subtypes of Antisocial Behavior questionnaire. We observed small but significant sex differences for a large number of socially aggressive behaviors, but these differences were largely a function of measurement non-invariance. These findings suggest that the inconsistencies found in prior studies examining sex differences in social aggression may be in part due to measurement non-invariance or similar forms of measurement bias.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Social aggression is defined as the use of social relationships to damage and inflict emotional harm on others (Cairns et al., 1989; Galen & Underwood, 1997), and includes behaviors that can be expressed overtly (e.g., threatening to end a friendship) or covertly (e.g., gossiping). The overall construct has been given various other names, including indirect and relational aggression (Feshbach 1969; Lagerspetz, Björkqvist, & Peltonen, 1988; Crick & Grotpeter, 1995). Although each term is defined slightly differently and may involve somewhat different behaviors (Xie, Cairns, & Cairns, 2002), prior research has indicated that all involve damaging the victim's social relationships or social status and the harmful manipulation of relationships (Card, Stucky, Sawalani, & Little, 2008; Archer & Coyne, 2005; Bjorkqvist, 2001; Underwood, Galen, & Paquette, 2001). Of these, the current study will focus primarily on social aggression.

Social aggression is quite damaging, both for the victims and the perpetrators. Available work has indicated that victims of social aggression often experience as much emotional distress as victims of physical aggression. Indeed, victims of social aggression are likely to experience emotional and social difficulties, including peer rejection, loneliness, psychological distress, internalizing difficulties, and problems with self-restraint (for a review, see Crick, Casas, & Nelson, 2002). The perpetrators of social aggression may also suffer from a number of maladaptive outcomes, although the evidence here is less consistent. Several studies have found evidence that social aggression is associated with peer rejection and poor quality friendships marked by conflict and instability (Crick & Grotpeter, 1995, Crick et al., 2006; Johnson & Foster, 2005; Ostrov & Crick, 2005; Kamper & Ostrov, 2013; Rose & Rudolph, 2006; Rose,

Swenson, & Carlson, 2004). Socially aggressive youth have also been found to have higher levels of loneliness, depression, and isolation than non-aggressive children (Crick & Grotpeter, 1995). Social aggression has also been associated with ADHD (Blachman & Hinshaw, 2002; Zalecki & Hinshaw, 2004), externalizing behavior (Keenan, Coyne, & Lahey, 2008; Spieker et al., 2012; Hipwell et al., 2002; Prinstein, Boergers, & Vernberg, 2001; Moretti, Holland, & McKay, 2001), borderline personality disorder features (Crick, Murray-Close, &Woods, 2005), and internalizing difficulties (Card et al., 2008). As noted above, however, the results to date have been somewhat inconsistent, with a small handful of studies either not replicating the above associations (Belden, Gaffrey, & Luby, 2012; Putallaz et al., 2007; Juliano, Werner, & Cassidy, 2006; Xie et al., 2002) or reporting that social aggression actually predicts higher social intelligence, popularity, and peer acceptance (Cillessen & Mayeux, 2004; Kaukiainen et al., 1999; Hart, Nelson, Robinson, Olsen, & McNeilly-Choque, 1998; Crick, Casas, & Mosher, 1997).

Importantly, however, social aggression may be especially damaging for and more predictive of problematic outcomes in females. Crick and colleagues (1995, 1997, 2006) found that socially aggressive girls reported higher levels of loneliness, isolation, and peer rejection than did socially aggressive boys and non-aggressive children. However, socially aggressive boys reported more depressive symptoms than did socially aggressive girls and non-aggressive children. Ostrov & Crick (2005) also found that socially aggressive girls engaged in less prosocial behavior than did socially aggressive boys or non-aggressive children. Collectively, these studies indicate that the relationship between social-psychological maladjustment and social aggression may vary as a function of sex, and that this association may be stronger or more pervasive for girls than for boys for some forms of psychosocial maladjustment.

Studies of sex differences in social aggression have not been confined to its external correlates, however. A large number of studies have examined whether females are more socially aggressive than males. According to popular gender stereotypes, social aggression is the "girl form" of aggression and physical aggression is the "boy form" of aggression. While previous research has definitively established that males are more physically aggressive than females, the investigation of sex differences in social aggression has yielded notably inconsistent results. Many studies have supported the popular notion that females are more socially aggressive than males (e.g. Bjorkqvist, Lagerspetz, & Kaukiainen, 1992; Crick & Grotpeter, 1995; Crick et al., 1997; Österman et al., 1998; Archer, 2004; Vaillancourt, 2005), while others found that boys engaged in higher levels of social aggression (e.g. Tomada & Schneider, 1997; Hennington, Hughes, Cavell, & Thompson, 1998; McEvoy, Estrem, Rodriguez, & Olson, 2003; Salmivalli & Kaukiainen, 2004; Moroschan, Hurd, & Nicoladis, 2009; Artz, Kassis, & Moldenhauer, 2013) or that there were no sex differences in prevalence (e.g. Hart et al., 1998; Deveaux & Daniels, 2000; Keenan et al., 2008; Belden et al., 2012; Kamper & Ostrov, 2013). These notable inconsistencies have led to a number of meta-analytic efforts to clarify the presence and direction of any sex differences. Such work has generally concluded that females do engage in slightly more socially aggressive behavior than do males, but that the magnitude of this difference is small (Archer, 2004; Card et al., 2008; Scheithaur, Haag, Mahlke, & Ittel, 2014). However, a more detailed investigation is warranted.

One intriguing possibility for these inconsistencies across sex relates to specific behaviors that comprise the construct of social aggression. Social aggression can be assessed using several different informants (including peer ratings, peer nominations, observation, self-report, teacher nominations, and parental ratings) and different measures (e.g., Direct and

Indirect Aggression Scale, Revised Social Experience Questionnaire, Aggression Questionnaire), each of which contains unique behavioral items (e.g., "tries to manipulate adults so that others get into trouble" is unique to the Revised Adolescent Social Experience Questionnaire). This lack of definitional and measurement clarity may contribute to the inconsistent sex differences observed in prior work (Pellegrini & Roseth, 2006). It also raises an important question: is it possible that the inconsistencies in sex differences reported at the level of the overall scale are related to the specific items? Put another way; are there consistent sex differences for specific behaviors?

To our knowledge, only one previous study has examined sex differences in the specific items. Coyne, Archer, and Eslea (2006) compared exposure to indirect, relational, social, verbal, and physical aggression in the school environment and on television in adolescents ages 11-15. Although there was no evidence for overall sex differences, differences were found in the frequency of exposure to some specific behaviors, such that girls reported observing more gossiping and dirty looks while boys reported observing more making fun of others and hitting in the school environment. Of note, however, Coyne and colleagues only examined exposure to aggressive behaviors and did not assess sex differences in the frequency with which adolescents engaged in these behaviors themselves.

In short, there is a clear need for a study that empirically evaluates the presence and direction of possible sex differences in the specific behaviors that comprise the construct of social aggression. The primary aim of the current study is to do just this, examining possible sex differences in individual socially aggressive behaviors. Our secondary aim was validate any sex differences we observed. We thus examined whether sex differences in specific social aggressive behaviors extended to their external correlates. Based on notable inconsistencies in

previous research, we do not have any predictions about the presence and/or direction of sex differences in specific socially aggressive behaviors.

METHODS

PARTICIPANTS

Four separate samples were examined as part of this study. Research protocols were approved by the Michigan State University IRB.

Samples 1 and 2. The first sample was an epidemiologic sample of 1000 children, nested in 500 twin pairs, who were assessed as part of the on-going Twin Study of Behavioral and Emotional Development in Children (TBED-C) within the Michigan State University Twin Registry (MSUTR) (Burt & Klump, 2013; Klump & Burt, 2006). The second sample was an independent, 'at-risk' sample of 1000 children, nested in 500 twin pairs, who were also assessed as part of the on-going TBED-C. To be eligible for participation in Sample 1, twins were required to live within two hours of our MSU laboratory, to be between the ages of 6 and 10 years-old at the time of recruitment, and could not have a cognitive or physical condition (e.g., significant developmental delays) that would preclude completion of the roughly 4-hour assessment (as assessed via parental report during the initial phone screen). The eligibility criteria for Sample 2 were the same as those for Sample 1, with the exception that families in Sample 2 were additionally required to live in moderately to severely disadvantaged neighborhoods (neighborhood poverty rates were required to be at or above the 2008 mean of 10.5% according to the US Census). Children gave informed assent, while parents gave informed consent for themselves and their children. Participating twins ranged from in age 6 to 10 years, although a small handful (n=16) had turned 11 by the time the family participated (mean age (SD) = 8.1 years (1.43)).

Department of Vital Records in the Michigan Department of Community Health identified twins in our age-range either directly from birth records or via the Michigan Twins

Project, a large-scale population-based registry of twins in lower Michigan that were themselves recruited via birth records. The Michigan Bureau of Integration, Information, and Planning Services database was used to locate family addresses within 120 miles of East Lansing, MI through parent drivers' license information. Pre-made recruitment packets were then mailed to parents on our behalf. A reply postcard was included for parents to indicate their interest in participating. Interested families were contacted directly by project staff. Parents who did not respond to the first mailing were sent additional mailings approximately one month apart until either a reply was received or up to four letters had been mailed.

This recruitment strategy yielded overall response rates of 62% for the population-based sample and 56% for the at-risk sample, which are similar to or better than those of other twin registries that use anonymous recruitment mailings (Baker, Barton, & Raine, 2002**;** Hay, McStephen, Levy, & Pearsall-Jones, 2002). Twins participating in our population-based study belonged to particular ethnic groups at rates comparable to area inhabitants (e.g., Caucasian non-Hispanic: 85.1% and 85.5%, African-American: 6.0% and 6.3% for the participating families and the local census, respectively). As expected, these proportions differed significantly from those in the at-risk sample (14.6% African-American and 76.9% Caucasian non-Hispanic; $X^2(1)=18.23$, $p < 0.001$). Importantly, both samples appear to be representative of recruited families (as assessed via a brief questionnaire screen administered to 80% of non-participating families). As compared to non-participating twins, participating twins were experiencing similar levels of conduct problems, emotional symptoms, or hyperactivity (Cohen's *d* effect sizes ranged from -0.08 to 0.01 in the population-based sample and 0.01 to 0.09 in the at-risk sample; all ns). Participating families in both samples also did not differ from non-participating families in paternal felony convictions (d = -0.01 and 0.13 for the population-based and the at-risk samples,

respectively), rate of single parent homes (d = 0.10 and -0.01 for the population-based and the at-risk samples, respectively), paternal years of education (both $d \leq 0.12$), or maternal and paternal alcohol problems (d ranged from 0.03 to 0.05 across the two samples). However, participating mothers in both samples reported slightly more years of education (d = 0.17 and 0.26, both $p <$ 0.05) than non-participating mothers. Maternal felony convictions differed across participating and non-participating families in the population-based sample (d = -0.20; $p < 0.05$) but not in the at-risk sample (d = .02). All told, we do not believe these differences significantly compromise the generalizability of these data.

Twin zygosity was determined using a standard 5-item questionnaire that assesses within-pair physical similarity and is over 95% accurate (Peeters, Gestel, Vlietinck, Derom, & Derom, 1998). One parent completed the questionnaire over the phone prior to the assessment. Unclear or discrepant zygosities are resolved by comparing twin sibling DNA markers (Klump & Burt, 2006). Half (50.3%) of the twins in sample 1 and one-third (33.3%) of the twins in sample 2 were monozygotic.

Samples 3 and 4. The third and fourth samples consisted of 700 and 500 undergraduate students, respectively, enrolled in psychology courses at a large, public university in the Midwest who participated in exchange for course credit or extra credit. The ethnic breakdown of Samples 3 and 4 was White (77.5%), Black (5.9%), Asian or Pacific Rim (8.0%), Hispanic (2.9%) and other (5.7%) ethnicities. Seventy (69.8%) of the samples were female (average age = 19.73 years; SD = 2.37). Research protocol was approved by the Michigan State University IRB. All participants will provide informed consent.

MEASURES

Social aggression. For all samples, socially aggressive behaviors were assessed using the Subtypes of Antisocial Behavior Questionnaire (STAB; Burt & Donnellan, 2009; Burt & Donnellan, 2010). The STAB is a 32-item measure assessing three major dimensions of antisocial behavior, one of which is social aggression. The Social Aggression Scale (SA) presents 11 behaviors (see Table 1 for full list of items) and asks the participants to report on the frequency with which they have committed each behavior. This scale ranges from 1 (never) to 5 (nearly all the time). Previous research estimates Cronbach's α to be between 0.83 to 0.90 for the Social Aggression Scale.

In samples 3 and 4, SA will be assessed via self-report. In samples 1 and 2, we collected data from the twins' mothers and teachers (due to concerns regarding the reliability and validity of self-reports of social aggression in young children). Numerous studies have indicated that multiple informants provide a more valid indication of a child's behavior that does any one informant alone (Bird, Gould, & Staghezza, 1992; Hart, Lahey, Loeber, & Hanson, 1994; Angold & Costello, 1996; Jensen et al., 1999; Hope et al., 1999). In addition to examining each informant report separately, SA items will be combined across informants using the "OR" approach. In this approach, the highest score on a given item from any informant is used as the score on that item. Items will then be summed to yield a multiple-informant scale score for each participant. This technique is often regarded as the most valid indication of a particular phenotype in children (Bird et al., 1992; Hart et al., 1994; Angold & Costello, 1996; Jensen et al., 1999; Hope et al., 1999). SA items will also be combined using averaging and "AND" (e.g. the lowest score on a given item from any informant is used as the score on that item) approaches to determine whether results are robust across pooling methods.

Adaptive and maladaptive functioning. The Achenbach System of Empirically Based Assessment (ASEBA) measures were used to assess adaptive and maladaptive functioning in the twins in Samples 1 and 2. The child's parent(s) completed the 113-item Child Behavior Checklist (CBCL) for each twin. Each twin's homeroom teacher completed a corresponding checklist, the Teacher Report Form (TRF). Additionally, each twin was interviewed using the Semistructured Clinical Interview for Children and Adolescents (SCICA). The CBCL, TRF, and SCICA all yield comparable scales. Importantly, these measures can be scored to produce both empirically based syndrome scales (e.g. Social Problems) and DSM-oriented scales (e.g. Attention Deficit/Hyperactivity Problems). Furthermore, these assessments assess functioning dimensionally, which allows for the evaluation of both adaptive and maladaptive functioning. These measures are validated for use in children ages 6-18 (Achenbach, 1991; Achenbach & Rescorla, 2001), show excellent reliability (Achenbach & Rescorla, 2001; McConaughy & Achenbach, 2001), and are widely used and well-known in child developmental research.

The college student assessments did not include measures of maladaptive functioning per se, but did include a measure of personality, which has been tied to multiple indices of adaptive and maladaptive functioning (Krueger, Caspi, & Moffitt, 2000; Samuel, Carroll, Rounsaville, & Ball, 2013; Suzuki, Samuel, Pahlen, & Krueger, 2015). The International Personality Item Pool—Five-Factor Model (Goldberg, 1999) was used to assess the personality traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Intellect/Imagination in the emergent adults in samples 3 and 4.

ANALYSES

Power analyses. Power analyses were conducted to determine the number of participants required to detect a mean effect size difference of .20 (a small effect). In the child samples, 960 families (e.g. 1920 twins) are needed to detect small effects with 90% statistical power using an alpha-level of .05. In the undergraduate sample, 1060 participants are necessary to detect small effects with 90% statistical power using an alpha-level of .05. The number of males and females in samples 1 and 2 are approximately even, but the sex distribution of samples 3 and 4 are less balanced. Generally, unbalanced designs require more participants than balanced designs to achieve an equivalent level of power. Therefore, samples 1 and 2 were combined for all child analyses and samples 3 and 4 were combined for all emerging adult analyses.

Sex differences. Linear regression was used to test whether the frequency of particular socially aggressive behaviors vary as a function of participant sex. Analyses in samples 1 and 2 were conducted via multi-level modeling, in which the child is the lower-level unit and the family is the upper-level unit, to control for the non-independence of the observations. The twin samples (Samples 1 and 2) were nested within their family unit. Multiple regression analyses was then conducted to evaluate whether sex moderates the relationship between socially aggressive behaviors and concurrent maladjustment in the child twin samples and personality in the emergent adult samples. Multilevel regression models, including hierarchical linear models, are an appropriate framework for research designs in which observations are nested or otherwise not independent.

Measurement invariance. The extent to which the Social Aggression Scale of the STAB exhibited measurement invariance between males and females was examined in order to determine whether the questionnaire measures identical constructs with the same structure across

the two groups. If measurement invariance holds, the Social Aggression Scale of the STAB can be reasonably used to compare the prevalence and consequences of social aggression in males and females. If measurement invariance does not hold, there may be limitations for comparing mean differences in the overall scale and individual behavior scores such that the mean differences found in the previous analyses may be due to artifactual—rather than true—differences between groups.

Following the step-by-step procedure outlined by van de Schoot, Lugtig, & Hox (2012), a set of increasingly constrained structural equation models were estimated using Mplus v.7.4 (Muthén & Muthén, 1998-2012) and then tested to determine whether differences between these models were significant. Robust maximum likelihood (MLR) estimation was used for all analyses and nested model comparisons were conducted using the Satorra-Bentler (TRd) scaled chi-square difference test. Females served as the reference group in all invariance models. Separate analyses were conducted for mother and teacher report of twin behavior and self-report of emergent adults.

First, a completely unconstrained, baseline configural invariance model was estimated across groups. In these models, factor loadings, intercepts, and residual variances were allowed to freely vary across groups. As shown in Tables 12-14, the configural models fit the data well in absolute terms (i.e., CFI >0.90 and RMSEA <0.08), so a series of model constraints were then applied in successive models to examine potential decreases in fit resulting from measurement non-invariance.

After the estimation of a baseline model, equality of the item factor loadings across groups was then examined in a full metric invariance model in which all factor loadings were constrained to be equal across groups, but all intercepts and residual variances were still

permitted to vary across groups. Next, equivalence of the item intercepts across groups was then examined in a full intercept invariance model in which all intercepts were constrained to be equal across groups, but all factor loadings and residual variances were still permitted to vary across groups. The potential change in fit of both the metric invariance and intercept invariance models were tested against the configural invariance baseline model. The next model to be tested was a full scalar invariance model in which all factor loadings and intercepts were constrained to be equal across groups, but residual variances were still permitted to vary across groups. The potential change in fit of the scalar invariance model was tested against the metric invariance model. The final model to be tested was a full uniqueness invariance model in which all factor loadings, intercepts, and residual variances were constrained to be equal across groups. The potential change in fit of the full uniqueness model was tested against the scalar invariance model.

When there was evidence of a significant drop in fit, tests of partial measurement invariance were conducted to determine which of the factor loadings and/or intercepts differed across groups. Items with the largest parameter discrepancies were identified as non-invariant and subsequently unconstrained across groups. When all non-invariant items were identified, a new, partially constrained invariance model was tested for potential decrease in model fit. When an item was found to be non-invariant in one model, then that scale's parameters were not constrained during any subsequent models.

Following these significance testing procedures, effect sizes were estimated to determine the practical significance of measurement non-invariance on mean-level differences in boys' and girls' social aggression scores. First, a latent trait estimation approach was used to determine the extent to which the presence of measurement non-invariance influenced observed trait-level

differences between groups (Tay et al., 2015). Put another way, this metric estimates the degree to which mean-level comparisons of social aggression across boys and girls were biased by measurement non-invariance. This overall effect size was calculated by subtracting the model-estimated latent trait mean of the fully constrained full uniqueness model from the model-estimated latent trait mean difference of the best-fitting partial uniqueness model. These effect sizes estimates are interpreted using the same guidelines as Cohen's *d*.

Next, an intercept differences approach was used to determine the extent to which the presence of non-invariance influenced observed item-level differences between groups (Millsap & Olivera-Aguilar, 2012). Specifically, this metric estimated the proportion of the observed mean-level differences in specific socially aggressive behaviors across boys and girls that is due to intercept non-invariance. This proportion was calculated by dividing the difference in intercepts as estimated by the partial scalar model by the difference in observed means for each non-invariant item. When an item is invariant, this value will be 0%, which indicates the observed mean difference for that item is solely due to true differences in levels of the latent trait of social aggression. However, values above 0% indicate the degree to which the observed mean difference is due to non-invariance rather than true latent trait differences.

RESULTS

INDIVIDUAL SOCIAL AGGRESSION ITEMS

Frequency

We first examined the frequency of individual socially aggressive behaviors (see Table 2). In middle childhood, "blamed others" was the most frequently reported socially aggressive behavior, regardless of informant or method of combining informant reports. Among college students, "gave someone the silent treatment when angry with him/her", "made fun of someone behind their back", and "blamed others" were reported most frequently. The least frequently reported behaviors in both middle childhood and college students were "intentionally damaged someone's reputation", "tried to turn others against someone when angry with him/her", and "revealed someone's secrets when angry with him/her".

Sex differences

We next evaluated whether there were mean sex differences in specific socially aggressive behaviors. To this end, multiple regression analyses were used to test whether the frequency of particular socially aggressive behaviors vary as a function of participant sex (see Table 3). Age and racial/ethnic background were added as covariates to all analyses. Separate analyses were conducted for each behavior.

In middle childhood, sex differences were not found for the overall scale of social aggression. However, sex differences were found for several specific behaviors. Girls were more likely than boys to give someone the silent treatment when angry (d = 0.12 to 0.21) and to reveal someone's secrets when angry (d = 0.11 to 0.14). By contrast, boys were more likely that girls to blame others (d = 0.15 to 0.27) and be rude towards others (d = 0.12 to 0.18). As a

sensitivity analysis, we evaluated whether these findings persisted once we controlled for familial-environmental background by restricting our analyses to the188 opposite-sex dizygotic twin pairs only. In the opposite-sex twin sample, sex differences were not found for the overall scale, but were again found for specific behaviors such that girls gave someone the silent treatment when angry more than boys (d = 0.21 to 0.27) and boys blamed others more than girls (d = 0.23 to 0.36). Mothers also reported that boys made fun of others (*d* = 0.16) and called others names behind their backs (*d* = 0.15) more than girls, but this finding was not robust across teacher-report and informant pooling methods. None of the other socially aggressive behaviors varied across sex in the middle childhood sample.

Among emergent adults, females reported engaging in significantly more social aggression overall (*d* = 0.17). At the item level, females also reported giving someone the silent treatment when angry (*d* = 0.52), revealing someone's secrets when angry (*d* = 0.09), calling someone names behind his/her back (*d* = 0.21), and making negative comments about another's appearance (*d* = 0.25) significantly more than males. By contrast, males reported that they were significantly more likely to intentionally damaged someone's reputation (*d* = 0.32) than were females. None of the other socially aggressive behaviors varied across sex in the college sample.

RELATIONSHIP BETWEEN SOCIALLY AGGRESSIVE BEHAVIORS AND EXTERNAL CORRELATES

The second aim of the study was to empirically evaluate the presence and direction of sex differences in the relationships between specific socially aggressive behaviors and external correlates. To this end, multiple regression analyses were conducted to evaluate whether sex moderated the relationship between socially aggressive behaviors and concurrent functioning and

clinical problems in the child twin samples, and to evaluate whether sex moderated the relationship between socially aggressive behaviors and personality, substance use, and other forms of aggression in the emergent adult samples. Age and racial/ethnic background were added as covariates to all analyses. Separate analyses were conducted for each behavior and the overall scale. Given the number of tests conducted, we primarily discuss only those results that were significant following Bonferroni correction.

Middle Childhood

Assessments of the twins' social aggression, functioning, and clinical problems were obtained from both the twins' mothers and their teachers. In the following regression analyses, the scores from one informant's report were used to predict the scores of the other informant's report (e.g. maternal assessment of social aggression was used to predict teacher assessment of functioning) in order to reduce the effects of informant bias on the results.

Competence/Adaptive Functioning

At the overall scale level, higher levels of teacher-reported social aggression predicted decreased school functioning and total functioning as measured by maternal report on the CBCL (see Table 4). These associations persisted to several items, and especially blamed others. Critically, however, there was no evidence that these associations varied across sex. Similarly, higher levels of maternal-reported social aggression predicted decreased appropriate behavior, happiness, and total functioning as measured by teacher report on the TRF (see Table 5). As before, these associations were not moderated by sex.

DSM-Oriented Scales

At the overall scale level, social aggression scores predicted increased oppositional defiant problems and conduct problems as measured by maternal report on the CBCL and teacher report on the TRF. Additionally, social aggression scores predicted increased attention deficit/hyperactivity problems as measured by maternal report, but not teacher report. Overall social aggression scores were not predictive of affective, anxiety, or somatic problems. With few exceptions, these patterns were found at the item-level as well (see Tables 6 and 7).

Child sex moderated the relationship between overall levels of social aggression and conduct problems as measured by maternal report, but not teacher report. According to maternal report, the relationship between social aggression and conduct problems is much stronger in girls than in boys. Moreover, this interaction persisted to two items: made fun other someone behind their back and tried to turn others against someone, even following Bonferroni correction. That said, it did not persist at either the item or the scale level when the informants were reversed (i.e., maternal reported social aggression and teacher-reported clinical problems), reducing confidence in the results.

Emergent Adulthood

Personality

At the overall scale level, social aggression scores predicted decreased agreeableness, decreased conscientiousness, and decreased emotional stability, and these relationships were not moderated by sex. Social aggression was not predictive of extraversion or intellect/imagination. This null pattern was generally found at the item-level as well (see Table 8).

Substance Use

At the overall scale level, social aggression scores predicted increased drug and alcohol use, and these relationships were not moderated by sex. This pattern was found at the item-level as well (see Table 9).

Aggression

At the overall scale level, social aggression scores predicted increased verbal aggression, increased physical aggression, increased anger, and increased hostility, and these relationships were not moderated by sex. With one exception (anger), these patterns were found at the item-level as well (see Table 10).

MEASUREMENT INVARIANCE

Finally, we used confirmatory factor analysis to test the extent to which the Social Aggression Scale of the STAB exhibited measurement invariance (i.e. equivalence) between males and females. Measurement invariance is a necessary prerequisite for mean-level comparisons between groups, and the presence of large amounts of measurement non-invariance (i.e. nonequivalence) may invalidate the previous findings regarding sex differences in socially aggressive behaviors. We first tested for metric invariance (also called weak or factorial invariance), which refers to whether factor loadings are equivalent across groups and tests whether factor indicators change at the same rate across groups with fluctuations in the latent factor (Millsap & Olivera-Aguilar, 2012). We then tested for intercept invariance, which refers to whether the factor intercepts are equal across groups and tests if scores on the factor indicators are the same across groups when the latent factor is at its zero-point (Millsap & Olivera-Aguilar,

2012). Finally, effect sizes were estimated to determine the impact non-invariance on mean-level differences in boys' and girls' social aggression scores. For ease of presentation, a summary of these analyses are found in Table 11. The individual fit statistics summarized in Table 11 are then presented in detail in Tables 12-14.

Middle Childhood

Maternal-Report

The full metric invariance model fit well and did not result in a significant decrease in fit relative to the configural model (Table 12; $\Delta X^2(10) = 12.525$, $p = 0.251$), indicating that factor loadings for maternal reports do not differ across child sex. However, the full intercept invariance model resulted in a significant decrease in fit relative to the configural model ($\Delta X^2(11) = 79.772$, $p < 0.001$). Non-variant intercepts were identified for "blamed others", "made fun of someone behind their back", "excluded someone from group activities when angry", "gave someone the silent treatment when angry", "call someone names behind back", and "was rude towards others".

Subsequent effect size analyses revealed that the overall effect of non-invariance on the latent trait of social aggression was 0.074, which represents a small bias effect, indicating that nearly all the observed sex differences for the social aggression total score was mostly due to true latent trait differences and not psychometric bias. In sharp contrast, however, non-invariance accounted for virtually all of the observed mean differences between girls' and boys' behavior scores for individual socially aggressive behaviors. Specifically, 97.9% of the mean differences found for "gave someone the silent treatment when angry" and 100% of the mean differences found for "blamed others", "made fun of someone behind their back", "call someone

names behind back", and "was rude towards others" was due to intercept non-invariance, indicating that the observed sex differences for these specific behaviors were due to psychometric bias rather than true differences in the latent trait of social aggression.

Teacher-Report

The full metric invariance model resulted in a significant decrease in fit relative to the configural model in teacher reports of twin social aggression (Table 13; $\Delta X^2(10) = 43.044$, $p < 0.001$). Non-variant factor loadings were identified for "blamed others", "tried to turn others against someone", and "was rude towards others", indicating that these specific behaviors were not equally related to the latent construct of social aggression in girls as compared to boys. The full intercept invariance model also resulted in a significant decrease in fit relative to the configural model ($\Delta X^2(11) = 182.450$, $p < 0.001$). Non-variant intercepts were identified for "blamed others", "excluded someone from group activities when angry", "gave someone the silent treatment when angry", "revealed someone's secrets when angry", "tried to turn others against someone", and "was rude towards others".

Subsequent effect size analyses revealed that the overall effect of non-invariance on the latent trait of social aggression was 0.001, which represents a negligible bias effect, indicating that any observed sex differences for the social aggression total score were due to true latent trait differences and not psychometric bias. As above, however, non-invariance accounted for 100% of the observed mean differences between girls' and boys' scores for "revealed someone's secrets when angry", indicating that the observed sex difference for this specific behavior was due to psychometric bias rather than true differences in the latent trait of social aggression. Unfortunately, we could not test the size of the effect of non-invariance for "blamed others",

"tried to turn others against someone" and "was rude towards others" because metric invariance was also violated for these specific behaviors.

Emergent Adulthood

The full metric invariance model resulted in a significant decrease in fit relative to the configural model in self-report of social aggression (Table 14; $\Delta X^2(10) = 20.974$, $p = 0.0213$). Non-variant factor loadings were identified for "intentionally damaged someone's reputation", indicating that this behavior was not equally related to the latent construct of social aggression in males and females. The full intercept invariance model also resulted in a significant decrease in fit relative to the configural model ($\Delta X^2(10) = 165.634$, $p < 0.001$). Non-variant intercepts were identified for "made fun of someone behind their back", "gave someone the silent treatment when angry", "intentionally damaged someone's reputation", "call someone names behind back", and "made negative comments about other's appearance".

Subsequent effect size analyses revealed that the overall effect of non-invariance on the latent trait of social aggression was 0.15, which represents a small bias effect, indicating that observed sex differences for the social aggression total score was mostly due to true latent trait differences and not psychometric bias. Again, however, non-invariance accounted for most of the observed mean differences in specific socially aggressive behaviors. Specifically, 98.5% of the mean differences found for "gave someone the silent treatment when angry", 99.4% of the mean differences found for "call someone names behind back", and 99.0% of the mean differences found for "made negative comments about other's appearance" was due to intercept non-invariance. However, none of the mean differences found for "revealed someone's secrets when angry" were due to non-invariance. We could not test the size of the effect of non-

invariance for “intentionally damaged someone’s reputation” because metric invariance was also violated for this behaviors. Such results indicate that observed sex differences for “revealed someone’s secrets when angry” were due to true differences in the latent trait of social aggression, but observed sex differences for the other behaviors were due to psychometric bias.

DISCUSSION

The primary aim of this study was to empirically evaluate the presence and direction of possible sex differences in the individual behaviors that comprise the construct of social aggression and to validate any sex differences we observed. This is the first study to investigate sex differences in the perpetration of specific socially aggressive behaviors and to test whether observed sex differences are due to measurement non-invariance. Sex differences were not found for the overall scale of social aggression in middle childhood, but were found for several specific behaviors. Girls gave someone the silent treatment when angry, revealed someone's secrets when angry, intentionally damaged someone's reputation, and tried to turn others against someone more than boys. Boys blamed others, were rude towards others, and made fun of someone behind their back more than girls. Sex differences in emergent adulthood were found for both the overall scale of social aggression and for several specific behaviors. Females gave someone the silent treatment when angry, revealed someone's secrets when angry, called someone names behind his/her back, and made negative comments about another's appearance significantly more than males. Males intentionally damaged someone's reputation more than females.

The observed pattern of sex differences initially imply that there may be consistent, genuine sex differences in engaging in some, but not all, socially aggressive behaviors. However, tests of measurement invariance at the item level revealed that, with the exception of "revealed someone's secrets when angry" in emergent adults, almost all of the observed sex differences in individual behaviors were due to intercept non-invariance. Consistent with this, none of the associations with external correlates differed in any meaningful way across sex. We thus conclude that observed sex differences in social aggression appear to represent measurement

issues, rather than legitimate differences across sex, and have few, if any, practical implications regarding the consequences of engaging in these behaviors.

Such results are critically important for our interpretation of mean-level sex differences because non-invariance implies that population differences in the observed means of these specific behaviors must be due to unknown influences other than the latent factor of social aggression. We can be confident that observed mean-level sex differences in only one behavior – revealed someone's secrets when angry in emergent adulthood – was the result of genuine differences in social aggression across sex, but the large effects of measurement non-invariance make it difficult to provide a straightforward interpretation of the observed differences in all other socially aggressive behaviors.

Although we are unable to investigate the causes of measurement non-invariance at this time, it is possible that popular gender stereotypes influenced informant report of child social aggression. For example, giving someone the silent treatment or revealing someone's secrets when angry may be perceived as covert, female behaviors rather than manifestations of direct, masculine aggressiveness, and therefore more likely to be reported on in females than in males. Another possibility is that some informants may have different "baselines" for the behavior of female and male children so that what counts as "a lot" of a behavior in girls might not count as "a lot" of the same behavior in boys. Future work should contrast these various possibilities in an effort to narrow in on the source(s) of measurement invariance.

Although measurement non-invariance largely influenced the observed sex differences at the item level, the impact of non-invariance on observed sex difference was minimal at the overall scale level. An effect size that is large at the individual-variable level may be washed out at the scale level if enough variables are included in the scale (Stark, Chernyshenko, & Drasgow,

2004). In other words, there are enough invariant behaviors in the scale that the overall effect of the noninvariant behaviors was minimized. Put differently, the overall score derived from Social Aggression Scale of the Subtypes of Antisocial Behavior questionnaire is largely invariant across sex in middle childhood and emergent adult samples and can confidently be used to make comparisons across groups as the same latent construct of social aggression is being assessed similarly in both groups. These results suggest that the previously-reported inconsistencies in observed sex differences may be due to ambiguous differences in questionnaire items rather than genuine differences in the underlying construct of social aggression. Failure to account for measurement non-invariance may have led to erroneous conclusions based on overestimated sex differences in the past. It will be important for future social aggression research to include tests of measurement invariance across targets and informants to ensure that the interpretations of observed group differences is due to the target latent variable and not more uncertain influences.

The second aim of the current study was to empirically evaluate the presence and direction of sex differences in the relationships between specific socially aggressive behaviors and external correlates. As discussed earlier, it is unclear whether engaging in social aggression is associated with adaptive or maladaptive functioning in perpetrators. Here, engaging in social aggression was consistently negatively associated with adaptive functioning and psychological health and positively associated with maladaptive functioning and psychological problems across samples, informants, and concurrent constructs. Socially aggressive individuals reported more problems at school and with peers, more externalizing symptoms, decreased emotional stability, and increased drug and alcohol use.

Some researchers have theorized that social aggression has a larger impact on girls as females place greater value on social relationships than do boys (Paquette & Underwood, 1999).

Other have concluded that social aggression has greater implications for boys because it is a sex-atypical form of aggression and, therefore, males who are highly socially aggressive are more maladjusted than their peers (Crick, 1997). Our results supported neither of these theories. Instead, we found that sex does not moderate the relationships between social aggression and concurrent social and psychological functioning, results that persisted across samples, ages, and informants. More precisely, we found that social aggression is neither a sex-typical nor sex-atypical form of aggression for males or females. Among emergent adults, females reported engaging in slightly more social aggression overall, but this difference had no impact on their adjustment. In short, social aggression may be more commonly associated with females in popular culture, but these associations do not stand up to empirical scrutiny.

The results of these analyses also have implications for the validity of various informant reports of social aggression in childhood. Pooling ratings or reports from multiple informants is often regarded as the most valid indication of a particular phenotype in children, but this method can prove to be expensive and time consuming for researchers. We found that although mothers rated their children as more socially aggressive than did teachers, a consistent pattern of results emerged across individual informants and data pooling methods for both mean-level sex differences and concurrent adjustment analyses. Maternal report of social aggression thus appear to be reasonable proxies for multi-informant reports.

Despite the strengths and depth of this study, it is important to consider some limitations. One limitation worth mentioning is that our investigation was limited to two distinct age groups, middle childhood and emergent adulthood, and it is thus possible that these results are specific to these development periods. In particular, the extent to which there are valid, mean-level sex differences in socially aggressive behaviors and the practical implications of these differences

should be examined in early-to-mid adolescence as social aggression typically peaks in the teenage years. Additionally, we focused exclusively on mother- and teacher-reports of social aggression in the children and self-reports in the emergent adults. Self- and peer-reports were not used to assess childhood social aggression as the validity of child-reports can be suspect due to developmental limitations on comprehension and responding or the influence of social desirability (Baker et al., 2007; Burt, McGue, Krueger, & Iacono, 2005). However, social aggression is characterized by not only overt behaviors that adults are likely to witness, but also covert behaviors of which only the child and his or her peers may be aware. Therefore, future research should incorporate self-, peer-, and adult-reports into the assessment of social aggression in childhood, adolescence, and emerging adulthood.

In conclusion, our study of sex differences in social aggression yielded several significant findings. The most salient was that we observed small but significant sex differences for a large number of specific socially aggressive behaviors, but these differences appear to be largely a function of measurement non-invariance. These findings suggest that the inconsistencies found in prior studies examining sex differences in social aggression may be in part due to measurement non-invariance or similar forms of measurement bias, such as structural non-invariance or differential item functioning. Moreover, these results illustrate the importance of taking a psychometrically informed approach to the study of group differences in social aggression and similar phenotypes. Without measurement equivalence, researchers may be drawing biased conclusions from observed mean-level group differences; therefore, item-level and scale-level effect sizes must be estimated to determine the practical impact of non-invariance.

APPENDIX

Table 1. Social Aggression Scale of the STAB

Item Number	Text Description
1	Blamed others
2	Tried to hurt someone's feelings
3	Made fun of someone behind their back
4	Excluded someone from group activities when angry with him/her
5	Gave someone the silent treatment when angry with him/her
6	Revealed someone's secrets when angry with him/her
7	Intentionally damaged someone's reputation
8	Tried to turn others against someone when angry with him/her
9	Call someone names behind his/her back
10	Was rude towards others
11	Made negative comments about other's appearance

Table 2. Frequency of perpetrating socially aggressive behaviors (means and standard deviations)

	Childhood				Young Adulthood	
	Maternal-Report		Teacher-Report		Self-Report	
	Boys	Girls	Boys	Girls	Males	Females
Blamed others	2.66 (0.91)	2.49 (0.91)	1.81 (0.99)	1.52 (0.85)	2.40 (0.77)	2.46 (0.76)
Tried to hurt someone's feelings	1.78 (0.79)	1.73 (0.74)	1.42 (0.69)	1.38 (0.72)	2.04 (0.78)	1.97 (0.76)
Made fun of someone behind their back	1.66 (0.74)	1.56 (0.67)	1.66 (0.74)	1.56 (0.67)	2.44 (0.80)	2.54 (0.78)
Excluded someone from group activities when angry	1.65 (0.78)	1.72 (0.80)	1.38 (0.65)	1.44 (0.76)	2.06 (0.88)	2.09 (0.88)
Gave someone the silent treatment when angry	1.83 (0.87)	1.92 (0.88)	1.38 (0.76)	1.44 (0.76)	2.40 (0.98)	2.91 (0.95)
Revealed someone's secrets when angry	1.30 (0.63)	1.31 (0.60)	1.07 (0.27)	1.12 (0.40)	1.34 (0.66)	1.40 (0.63)
Intentionally damaged someone's reputation	1.01 (0.12)	1.02 (0.15)	1.01 (0.14)	1.02 (0.20)	1.26 (0.55)	1.11 (0.36)
Tried to turn others against someone	1.26 (0.55)	1.27 (0.56)	1.11 (0.37)	1.20 (0.54)	1.53 (0.76)	1.50 (0.73)
Call someone names behind back	1.57 (0.71)	1.47 (0.64)	1.28 (0.57)	1.30 (0.61)	2.07 (0.83)	2.24 (0.78)
Was rude towards others	1.82 (0.77)	1.74 (0.73)	1.45 (0.77)	1.36 (0.73)	2.18 (0.80)	2.24 (0.72)
Made negative comments about other's appearance	1.59 (0.71)	1.60 (0.69)	1.18 (0.45)	1.19 (0.49)	1.99 (0.83)	2.19 (0.79)
Overall Scale	18.14 (4.99)	17.83 (4.84)	14.52 (4.70)	14.38 (5.48)	21.71 (5.65)	22.66 (5.43)

Note: The Social Aggression Scale asks the participants to report on the frequency with which they have committed each behavior. This scale ranges from 1 (never) to 5 (nearly all the time).

Table 3. Summary of sex differences in perpetrating socially aggressive behaviors

	Middle Childhood					Emergent Adulthood
	Maternal-Report	Teacher-Report	Combined: Average	Combined: OR approach	Combined: AND approach	Self-Report
Blamed others	Males**	Males**	Males**	Males**	Males**	-
Tried to hurt someone's feelings	-	-	-	-	-	-
Made fun of someone behind their back	Males*	-	-	Males*	-	-
Excluded someone from group activities when angry	-	-	-	-	-	-
Gave someone the silent treatment when angry	Females*	-	Females*	Females*	Females*	Females**
Revealed someone's secrets when angry	-	Females*	Females*	-	-	Females*
Intentionally damaged someone's reputation	-	-	-	Females*	-	Males**
Tried to turn others against someone	-	Females**	Females*	Females*	-	-
Call someone names behind back	Males*	-	-	-	-	Females**
Was rude towards others	Males*	Males*	Males*	Males*	Males*	-
Made negative comments about other's appearance	-	-	-	-	-	Females**
Overall Scale	-	-	-	-	-	Females*

Note: * and ** indicate that frequency differs between sex at $p \leq .05$ and .0001 (Bonferroni corrected), respectively. Cohen's *d* effect sizes are reported in the text. Age and ethnicity were added as covariates to all analyses.

Table 4. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors (teacher report) and adaptive functioning (maternal report) in middle childhood

	Activities	School	Social	Total
	ME/INT	ME/INT	ME/INT	ME/INT
Blamed others	-.065/.004	-.134**/-.007	-.186*/.001	-.388**/.005
Tried to hurt someone's feelings	-.021/.094	-.141**/.002	-.134*/-.041	-.275*/.071
Made fun of someone behind their back	.020/.002	-.162**/.016	-.080/.025	-.247*/.050
Excluded someone from group activities	-.047/.054	-.120*/.001	-.222*/-.084	-.374*/-.015
Gave someone the silent treatment when angry	-.032/-.062	-.097*/.027	-.177*/-.024	-.318*/-.057
Revealed someone's secrets when angry	.078/-.124	-.178*/-.054	.020/-.168	-.024/-.393
Intentionally damaged someone's reputation	-.053/-.018	-.138/.130	-.243/.167	-.375/.250
Tried to turn others against someone	-.030/-.032	-.136*/-.068	-.190*/-.117	-.310/-.225
Call someone names behind back	-.074/.085	-.189**/-.016	-.199*/.026	-.445*/.126
Was rude towards others	-.002/.021	-.171**/-.030	-.165*/-.066	-.323*/-.065
Made negative comments about other's appearance	-.066/-.047	-.229**/-.092	-.167/-.031	-.439*/-.168
Total Score	-.004/.002	-.027**/-.005	-.032*/-.010	-.063**/-.012

Note: ME indicates the main effect of the socially aggressive behavior on the outcome. INT indicates the interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0001 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses.

Table 5. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors (maternal report) and adaptive functioning (teacher report) in middle childhood

	Academic	Working Hard	Behaving Appropriately	Learning	Happy	Total
	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT
Blamed others	.003/.014	-.092*/.058	-.153**/.026	-.006/.033	.005*/.039	-.352*/.157
Tried to hurt someone's feelings	.028/ -.009	-.088/.013	-.112*/.071	-.024/ -.013	-.125*/.042	-.333*/.128
Made fun of someone behind their back	.050/.000	-.145*/-.039	-.168*/.018	.025/-.069	-.118*/-.083	-.383*/-.172
Excluded someone from group activities	.026/.001	-.065/.038	-.121*/.026	.040/-.013	-.131*/-.029	-.272/.033
Gave someone the silent treatment when angry	.001/ -.015	-.054/-.006	-.063/.028	.013/-.039	-.094*/-.032	-.197/-.053
Revealed someone's secrets when angry	-.001/ .037	.030/.069	-.030/.071	.032/.090	-.025/-.029	.033/.213
Intentionally damaged someone's reputation	-.107/.138	-.330/.282	-.393/.417	.058/.283	-.227/.359	-.812/1.499*
Tried to turn others against someone	-.062/ -.028	-.234*/.043	-.278**/.034	-.114/.000	-.199*/.003	-.836**/.055
Call someone names behind back	-.028/.001	-.166*/.013	-.252**/.066	-.103 /.020	-.180**/.010	-.686**/.111
Was rude towards others	.025/.005	-.112*/.004	-.237**/.012	.019/-.013	-.105*/.058	-.423*/.053
Made negative comments about other's appearance	.049/.004	-.075/.002	-.206*/.035	.036/.001	-.096/-.021	-.335/.027
Total Score	.003/.001	-.024*/.005	-.039**/.009	-.001/ -.001	-.029**/ .000	-.094**/ .015

Note: ME indicates a significant main effect of the socially aggressive behavior on the outcome. INT indicates a significant interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0001 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses.

Table 6. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors (teacher report) and clinical problems (maternal report) in middle childhood

	Affective	Anxiety	Somatic	AD/H	Oppositional Defiant	Conduct
	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT
Blamed others	.152*/.013	.136*/.096	-.043/-.037	.693**/.088	.410**/-.047	.511**/.067
Tried to hurt someone's feelings	.084/.122	.085/.164*	-.044/-.066	.770**/.218	.579**/.204*	.747**/.239*
Made fun of someone behind their back	-.006/.059	-.003/.096	-.044/.039	.603**/.106	.431**/.227*	.675**/.356**
Excluded someone from group activities	.089/-.029	-.006/.041	-.031/-.003	.655**/.168	.401**/.080	.600**/.235*
Gave someone the silent treatment when angry	.177*/.082	.053/.089	-.040/-.012	.349*/.004	.263**/.051	.498**/.019
Revealed someone's secrets when angry	-.289*/.063	-.216/ -.008	-.230*/ .122	.534*/.185	.405*/.142	.842**/.395*
Intentionally damaged someone's reputation	-.014/.480	-.166/ -.118	-.102/-.017	-.445/-.684	.174/.442	.392/-.644
Tried to turn others against someone	.290*/.293*	.146/.204	-.117/-.074	1.008**/ .590*	.585**/.411*	1.16**/.747**
Call someone names behind back	.046/.065	.086/.030	-.038/.051	.643**/-.066	.542**/.010	.748**/.237*
Was rude towards others	.200*/.068	.156*/.086	-.071/-.097	.729**/.022	.569**/.136	.740**/.158
Made negative comments about other's appearance	.211*/.141	.098/.270*	-.051/.031	.765**/.109	.584**/.270*	.857**/.358*
Total Score	.022*/.018	.016/.021*	-.010/-.003	.117**/.032	.082**/.029*	.126**/.049**

Note: ME indicates a significant main effect of the socially aggressive behavior on the outcome. INT indicates a significant interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0001 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses.

Table 7. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors (maternal report) and clinical problems (teacher report) in middle childhood

	Affective	Anxiety	Somatic	AD/H	Oppositional Defiant	Conduct
	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT
Blamed others	.171*/-.066	.029/-.030	.021/-.033	.461*/.118	.209**/-.047	.363**/.030
Tried to hurt someone's feelings	.120*/.011	.035/.052	-.017/.017	.209/-.170	.213**/-.009	.309**/.020
Made fun of someone behind their back	.112/.045	-.029/-.012	.024/.019	.284/.096	.204*/.081	.302*/.025
Excluded someone from group activities	.074/.053	.043/.012	-.016/-.013	.221/.273	.197**/-.016	.264*/.068
Gave someone the silent treatment when angry	.051/-.031	-.051/.015	-.014/-.003	.018/-.017	.071/.001	.126/.022
Revealed someone's secrets when angry	-.078/.013	-.014/-.013	-.002/-.002	-.092/-.040	.055/.040	.172/.073
Intentionally damaged someone's reputation	-.004/-.242	-.115/-.235	-.121/.053	.471/-.421	.713*/-.485	.538/-.894*
Tried to turn others against someone	.184*/-.004	-.005/-.048	.034/-.004	.663*/.417	.321**/.053	.512**/.115
Call someone names behind back	.127/.022	-.015/-.029	.030/-.040	.560*/.068	.298**/.072	.406**/.029
Was rude towards others	.119/.045	.019/-.011	-.003/.018	.590*/.318	.303**/.133*	.336**/.168*
Made negative comments about other's appearance	.087/.026	-.035/-.023	.041/-.027	.098/.131	.230**/.021	.328**/.013
Total Score	.025*/.001	.000/-.001	.002/-.002	.074*/.027	.051**/.006	.075**/.012

Note: ME indicates a significant main effect of the socially aggressive behavior on the outcome. INT indicates a significant interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0001 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses.

Table 8. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors and personality traits in emergent adulthood

	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Intellect/Imagination
	ME/INT	ME/INT	ME/INT	ME/INT	ME/INT
Blamed others	-.024/-.013	-.043/.036	-.070*/.035	-.247**/.009	.023/.033
Tried to hurt someone's feelings	-.002/.004	-.127**/.018	-.085*/.044	-.258**/-.002	.034/-.008
Made fun of someone behind their back	.049/-.010	-.041/.035	-.042/.060*	-.175**/.002	.044/.045
Excluded someone from group activities when angry	.051/-.002	-.102**/-.011	-.077*/.013	-.180**/.030	.025/.035
Gave someone the silent treatment when angry	-.012/.028	.025/.052	-.028/.016	-.252**/-.015	.060, .080*
Revealed someone's secrets when angry	.027/-.043	-.114**/-.019	-.115**/-.002	-.174**/.006	-.066*/.028
Intentionally damaged someone's reputation	.044/-.022	-.113**/.015	-.108**/-.002	-.150**/.033	-.011/-.012
Tried to turn others against someone	.042/-.028	-.113**/-.024	-.104**/-.023	-.231**/-.004	-.006/.029
Call someone names behind back	.065*/-.026	-.084**/-.014	-.047/.074*	-.218**/.018	-.008/.034
Was rude towards others	-.007/-.031	-.119**/.018	-.073*/.048	-.237**/.021	.017/.063*
Made negative comments about other's appearance	-.029/-.065	-.116**/.035	-.185**/-.060	-.133**/.012	-.110**/.007
Total Score	.037/-.014	-.123**/.013	-.104**/.039	-.325**/.008	.024/.057

Note: ME indicates a significant main effect of the socially aggressive behavior on the outcome. INT indicates a significant interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0003 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses.

Table 9. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors and substance use in emergent adulthood

	Alcohol Use	Drug Use
	ME/INT	ME/INT
Blamed others	.033/-.091*	.085*/-.033
Tried to hurt someone's feelings	.120**/-.029	.207**/.004
Made fun of someone behind their back	.180**/-.061	.120*/-.017
Excluded someone from group activities	.216**/.045	.180**/.006
Gave someone the silent treatment when angry	.143**/-.035	.133**/-.033
Revealed someone's secrets when angry	.121**/-.063	.135**/-.082*
Intentionally damaged someone's reputation	.090*/-.027	.189**/-.036
Tried to turn others against someone	.196**/.008	.220**/.084*
Call someone names behind back	.159**/-.061	.159**/.003
Was rude towards others	.122**/-.089*	.130**/-.058
Made negative comments about other's appearance	.085*/.039	.210**/-.008
Total Score	.228**/-.046	.246**/-.002

Note: ME indicates a significant main effect of the socially aggressive behavior on the outcome. INT indicates a significant interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0003 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses

Table 10. Summary of results from multiple regression analyses examining whether sex moderates the relationship between socially aggressive behaviors and other aggressive constructs in emergent adulthood

	Verbal Aggression	Physical Aggression	Anger	Hostility
	ME/INT	ME/INT	ME/INT	ME/INT
Blamed others	.185**/.020	.130**/ .002	.185**/-.091**	.226**/ -.020
Tried to hurt someone's feelings	.293**/ .033	.201**/ .036	.232**/-.066*	.250**/ -.002
Made fun of someone behind their back	.188**/ .014	.086**/ .027	.146**/ -.048	.232**/ -.014
Excluded someone from group activities	.237**/ -.010	.187**/ .015	.217**/-.071*	.241**/ -.054
Gave someone the silent treatment when angry	.214**/ .000	.135**/ -.012	.226**/ -.042	.247**/ .004
Revealed someone's secrets when angry	.144**/ -.054	.116**/ -.014	.175**/ -.060*	.211**/ -.018
Intentionally damaged someone's reputation	.191**/ -.035	.149**/ -.029	.199**/-.089**	.187**/ -.048
Tried to turn others against someone	.223**/ .008	.204**/ .041	.259**/ -.021	.280**/ -.011
Call someone names behind back	.231**/ .014	.181**/ .024	.198**/ -.049	.271**/ -.014
Was rude towards others	.266**/ .017	.155**/ .001	.234**/-.060*	.231**/ -.029
Made negative comments about other's appearance	.056*/ -.049	.117**/ -.029	.155**/ -.058	.120**/ -.071
Total Score	.336**/ .017	.231**/ .023	.311**/-.079*	.361**/ -.022

Note: ME indicates a significant main effect of the socially aggressive behavior on the outcome. INT indicates a significant interaction between the socially aggressive behavior and sex on the outcome. * and ** indicate that frequency differs between sex at $p \leq .05$ and .0003 (Bonferroni corrected), respectively. Age and ethnicity were added as covariates to all analyses

Table 11. Summary of measurement non-invariant parameters of the STAB Social Aggression Scale (i.e. those that vary across sex)

	Factor Loadings			Intercepts		
	Childhood Maternal-Report	Childhood Teacher-Report	Emergent Adult Self-Report	Childhood Maternal-Report	Childhood Teacher-Report	Emergent Adult Self-Report
Blamed others	-	X	-	X	X	-
Tried to hurt someone's feelings	-	-	-	-	-	-
Made fun of someone behind their back	-	-	-	X	-	X
Excluded someone from group activities when angry	-	-	-	X	X	-
Gave someone the silent treatment when angry	-	-	-	X	X	X
Revealed someone's secrets when angry	-	-	-	-	X	-
Intentionally damaged someone's reputation	-	-	X	-	-	X
Tried to turn others against someone	-	X	-	-	X	-
Call someone names behind back	-	-	-	X	-	X
Was rude towards others	-	X	-	X	X	-
Made negative comments about other's appearance	-	-	-	-	-	X

Note: X indicates that the item violated measurement invariance across sex. Non-invariant factor loadings indicate that the factor loadings are not equivalent across groups and that factor indicators do not change at the same rate across groups with fluctuations in the latent factor of social aggression. Non-invariant intercepts indicate that the factor intercepts are not equal across groups. Effect sizes for intercept non-invariance and overall non-invariance are reported in the text.

Table 12. Model fit statistics for tests of multiple group measurement invariance in maternal reports of middle childhood social aggression

	X^2	df	Scaling Correction Factor	ΔX^2	Δdf	p	CFI	RMSEA
BASELINE CONFIGURAL	558.757	88	1.2995	-	-	-	0.904	0.074
FULL METRIC	546.623	98	1.3751	12.5254	10	0.251455	0.909	0.068
Blamed others	543.775	97	1.3785	10.9204	9	0.281232	0.909	0.069
Tried to hurt someone's feelings	542.848	97	1.3797	10.5656	9	0.306707	0.909	0.068
Made fun of someone behind their back	545.245	97	1.3750	12.0939	9	0.208117	0.909	0.068
Excluded someone from group activities when angry	544.278	97	1.3771	10.9654	9	0.278116	0.909	0.069
Gave someone the silent treatment when angry	544.819	97	1.3760	11.0952	9	0.269253	0.909	0.069
Revealed someone's secrets when angry	549.813	97	1.3669	12.5546	9	0.183847	0.908	0.069
Intentionally damaged someone's reputation	571.014	97	1.3056	14.2181	9	0.114779	0.904	0.071
Tried to turn others against someone	546.680	97	1.3704	11.1771	9	0.263771	0.909	0.069
Call someone names behind back	539.986	97	1.3795	8.6995	9	0.465463	0.910	0.068
Was rude towards others	545.428	97	1.3780	11.8827	9	0.220044	0.909	0.069
Made negative comments about other's appearance	546.274	97	1.3751	11.8605	9	0.221317	0.909	0.069
FULL INTERCEPT	636.240	99	1.2631	79.7716	11	< 0.00001	0.891	0.074
Blamed others	574.011	89	1.2961	17.9265	1	0.000230	0.901	0.075
Tried to hurt someone's feelings	561.930	89	1.2961	2.2196	1	0.136268	0.904	0.074
Made fun of someone behind their back	566.225	89	1.2960	7.8167	1	0.005177	0.903	0.074
Excluded someone from group activities when angry	563.470	89	1.2961	4.2218	1	0.039908	0.903	0.074
Gave someone the silent treatment when angry	564.735	89	1.2961	5.8665	1	0.015432	0.903	0.074
Revealed someone's secrets when angry	560.216	89	1.2961	0.0088	1	0.925261	0.904	0.074
Intentionally damaged someone's reputation	562.229	89	1.2959	2.5409	1	0.110932	0.904	0.074

Table 12 (cont'd)

Tried to turn others against someone	560.458	89	1.2961	0.3058	1	0.580269	0.904	0.074
Call someone names behind back	566.849	89	1.2960	8.6352	1	0.003297	0.903	0.074
Was rude towards others	565.402	89	1.2961	6.7337	1	0.009461	0.903	0.074
Made negative comments about other's appearance	560.542	89	1.2961	0.4151	1	0.519392	0.904	0.074

Note: Models that resulted in a significant reduction in model fit are highlighted in bold text. Bolded behaviors in the metric invariance model are those in which the factor loadings were not equivalent across groups. Bolded behaviors in the intercept invariance model are those in which the factor intercepts were not equal across groups. Scalar models test whether both factor loadings and intercepts can constrained to be equal across groups. Uniqueness models test whether factor loadings, intercepts, and residual variances can be constrained to be equal across groups. Partial models are those in which the parameters of the non-invariant items were unconstrained across groups.

Table 13. Model fit statistics for tests of multiple group measurement invariance in teacher reports of middle childhood social aggression

	X^2	df	Scaling Correction Factor	ΔX^2	Δdf	p	CFI	RMSEA
BASELINE CONFIGURAL	373.042	88	2.4176	-	-	-	0.925	0.068
FULL METRIC	416.524	98	2.6164	43.0449	10	< 0.00001	0.916	0.068
Blamed others	406.308	97	2.6309	17.2226	1	0.000033	0.918	0.067
Tried to hurt someone's feelings	414.480	97	2.6259	0.8321	1	0.361666	0.916	0.068
Made fun of someone behind their back	412.904	97	2.6271	3.2013	1	0.07358	0.917	0.068
Excluded someone from group activities when angry	413.524	97	2.6217	2.6911	1	0.10091	0.917	0.068
Gave someone the silent treatment when angry	415.669	97	2.6139	1.1460	1	0.284388	0.916	0.068
Revealed someone's secrets when angry	417.585	97	2.5718	2.2828	1	0.130816	0.915	0.068
Intentionally damaged someone's reputation	422.513	97	2.4802	2.6458	1	0.103824	0.914	0.069
Tried to turn others against someone	392.791	97	2.5815	12.6303	1	0.00038	0.922	0.066
Call someone names behind back	413.268	97	2.6179	3.1968	1	0.073783	0.917	0.068
Was rude towards others	409.080	97	2.6226	8.4070	1	0.003738	0.918	0.067
Made negative comments about other's appearance	417.251	97	2.6105	0.1755	1	0.675269	0.916	0.068
FULL INTERCEPT	455.987	99	2.2370	182.4501	11	< 0.00001	0.906	0.071
Blamed others	388.230	89	2.4021	29.5742	1	< 0.00001	0.921	0.069
Tried to hurt someone's feelings	376.106	89	2.4022	1.5430	1	0.214172	0.924	0.067
Made fun of someone behind their back	375.878	89	2.4020	0.9645	1	0.326056	0.924	0.067
Excluded someone from group activities when angry	376.197	89	2.4013	1.5467	1	0.213623	0.924	0.067

Table 13 (cont'd)

Gave someone the silent treatment when angry	376.584	89	2.4018	2.5837	1	0.107969	0.924	0.067
Revealed someone's secrets when angry	378.076	89	2.4002	6.4346	1	0.011192	0.924	0.068
Intentionally damaged someone's reputation	376.264	89	2.4002	1.4298	1	0.231797	0.924	0.067
Tried to turn others against someone	380.566	89	2.4004	13.1307	1	0.000291	0.923	0.068
Call someone names behind back	375.556	89	2.4016	0.0694	1	0.792212	0.924	0.067
Was rude towards others	378.217	89	2.4023	6.3684	1	0.011617	0.924	0.068
Made negative comments about other's appearance	375.531	89	2.4017	0.0463	1	0.829631	0.924	0.067

Note: Models that resulted in a significant reduction in model fit are highlighted in bold text. Bolded behaviors in the metric invariance model are those in which the factor loadings were not equivalent across groups. Bolded behaviors in the intercept invariance model are those in which the factor intercepts were not equal across groups. Scalar models test whether both factor loadings and intercepts can constrained to be equal across groups. Uniqueness models test whether factor loadings, intercepts, and residual variances can be constrained to be equal across groups. Partial models are those in which the parameters of the non-invariant items were unconstrained across groups.

Table 14. Model fit statistics for tests of multiple group measurement invariance in self-reported social aggression in emergent adults

	X^2	df	Scaling Correction Factor	ΔX^2	Δdf	p	CFI	RMSEA
BASELINE CONFIGURAL	464.003	88	1.2013	-	-	-	0.906	0.079
FULL METRIC	484.954	98	1.2014	20.9743	10	0.021276	0.903	0.076
Blamed others	482.147	97	1.2005	17.9518	9	0.035746	0.903	0.076
Tried to hurt someone's feelings	481.915	97	1.2044	18.6373	9	0.028463	0.904	0.076
Made fun of someone behind their back	484.078	97	1.2014	20.0973	9	0.017323	0.903	0.076
Excluded someone from group activities when angry	483.402	97	1.2047	20.1524	9	0.016997	0.903	0.076
Gave someone the silent treatment when angry	484.057	97	1.2036	20.5566	9	0.014775	0.903	0.076
Revealed someone's secrets when angry	485.659	97	1.1984	21.0308	9	0.012518	0.903	0.077
Intentionally damaged someone's reputation	472.673	97	1.2008	8.5114	9	0.483540	0.906	0.075
Tried to turn others against someone	483.783	97	1.2013	19.7800	9	0.019320	0.903	0.076
Call someone names behind back	484.205	97	1.2023	20.4218	9	0.015485	0.903	0.058
Was rude towards others	484.349	97	1.2026	20.6295	9	0.014404	0.903	0.076
Made negative comments about other's appearance	484.031	97	1.2037	20.5525	9	0.014795	0.903	0.076
FULL INTERCEPT	616.055	98	1.1851	165.6339	10	< 0.00001	0.870	0.088
Blamed others	466.858	89	1.1987	2.2846	1	0.130664	0.905	0.079
Tried to hurt someone's feelings	466.788	89	1.1987	2.1981	1	0.138181	0.905	0.079
Made fun of someone behind their back	468.545	89	1.1992	4.4089	1	0.035752	0.905	0.079
Excluded someone from group activities when angry	465.226	89	1.1993	0.5265	1	0.468082	0.906	0.079
Gave someone the silent treatment when angry	527.310	89	1.1998	70.4811	1	< 0.00001	0.890	0.085
Revealed someone's secrets when angry	466.054	89	1.2012	2.0272	1	0.154505	0.905	0.079

Table 14 (cont'd)

Intentionally damaged someone's reputation	487.988	89	1.1931	52.6229	1	< 0.00001	0.900	0.081
Tried to turn others against someone	465.710	89	1.1984	0.7422	1	0.388957	0.906	0.079
Call someone names behind back	474.021	89	1.2004	10.3532	1	0.001293	0.903	0.080
Was rude towards others	465.461	89	1.1992	0.7630	1	0.382391	0.906	0.079
Made negative comments about other's appearance	478.138	89	1.2007	14.5426	1	0.000137	0.902	0.080

Note: Models that resulted in a significant reduction in model fit are highlighted in bold text. Bolded behaviors in the metric invariance model are those in which the factor loadings were not equivalent across groups. Bolded behaviors in the intercept invariance model are those in which the factor intercepts were not equal across groups. Scalar models test whether both factor loadings and intercepts can constrained to be equal across groups. Uniqueness models test whether factor loadings, intercepts, and residual variances can be constrained to be equal across groups. Partial models are those in which the parameters of the non-invariant items were unconstrained across groups.

REFERENCES

REFERENCES

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Angold, A. & Costello, E.J. (1996). The relative diagnostic utility of child and parent reports of oppositional defiant behaviors. *International Journal of Methods in Psychiatric Research*, 6, 253-259.

Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology, 8*(4), 291-322.

Archer, J., & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review, 9*(3), 212-230.

Artz, S., Kassis, W., & Moldenhauer, S. (2013). Rethinking indirect aggression: The end of the mean girl myth. *Victims & Offenders, 8*(3), 308-328.

Baker, L.A., Barton, M., & Raine, A. (2002) The Southern California Twin Register at the University of Southern California. *Twin Research*, 5, 456–459.

Baker, L. A., Jacobson, K. C., Raine, A., Lozano, D. I., & Bezdjian, S. (2007). Genetic and environmental bases of childhood antisocial behavior: A multi-informant twin study. Journal of *Abnormal Psychology, 116*, 219–235.

Belden, A. C., Gaffrey, M. S., & Luby, J. L. (2012). Relational aggression in children with preschool-onset psychiatric disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(9), 889-901.

Bird, H.R., Gould, M., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child & Adolescent Psychiatry*, 31, 78-85

Björkqvist, K., Lagerspetz, K. M., & Kaukiainen, A. (1992). Do girls manipulate and boys fight? developmental trends in regard to direct and indirect aggression. *Aggressive Behavior, 18*(2), 117-127.

Björkqvist, K. (2001). Different names, same issue. *Social Development, 10*(2), 272-274.

Blachman, D. R., & Hinshaw, S. P. (2002). Patterns of friendship among girls with and without

attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology*, *30*(6), 625-640.

Burt, S.A., & Donnellan, M.B. (2009). Development and validation of the subtypes of antisocial behavior questionnaire. *Aggressive Behavior*, 35(5), 376-398.

Burt, S.A. & Donnellan, M.B. (2010). Evidence that the Sub-Types of Antisocial Behavior questionnaire (STAB) predicts momentary reports of acting-out behaviors. *Personality and Individual Differences*, 48, 917- 920.

Burt, S. A., & Klump, K. L. (2013). The Michigan State University Twin Registry (MSUTR): An update. *Twin Research and Human Genetics, 16*(1), 344-350.

Burt, S. A., McGue, M., Krueger, R. F., & Iacono, W. G. (2005). Sources of covariation among the child-externalizing disorders: Informant effects and the shared environment. *Psychological Medicine, 35*, 1133–1144.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Ferguson, L. L., & Gariépy, J. (1989). Growth and aggression: I. childhood to early adolescence. *Developmental Psychology, 25*(2), 320-330.

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. Child Development, 79(5), 1185-1229.

Cillessen, A. H. N., & Mayeux, L. (2004). From censure to reinforcement: Developmental changes in the association between aggression and social status. Child Development, 75(1), 147-163.

Coyne, S. M., Archer, J., & Eslea, M. (2006). "We're not friends anymore! unless...": The frequency and harmfulness of indirect, relational, and social aggression. *Aggressive Behavior, 32*(4), 294-307.

Crick, N.R. (1997) Engagement in gender normative versus nonnormative forms of aggression: Links to social–psychological adjustment. *Developmental Psychology, 33*(4), 610-617.

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66*(3), 710-722.

Crick, N. R., Casas, J. F., & Mosher, M. (1997). Relational and overt aggression in preschool. *Developmental Psychology, 33*(4), 579-588.

Crick, N. R., Casas, J. F., & Nelson, D. A. (2002). Toward a more comprehensive understanding of peer maltreatment: Studies of relational victimization. *Current Directions in Psychological Science, 11*(3), 98-101.

Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: A short-term longitudinal study. *Development and Psychopathology, 17*(4), 1051-1070.

Crick, N. R., Ostrov, J. M., Burr, J. E., Cullerton-Sen, C., Jansen-Yeh, E., & Ralston, P. (2006). A longitudinal study of relational and physical aggression in preschool. *Journal of Applied Developmental Psychology, 27*(3), 254-268.

Delveaux, K. D., & Daniels, T. (2000). Children's social cognitions: Physically and relationally aggressive strategies and children's goals in peer conflict situations. *Merrill-Palmer Quarterly, 46*(4), 672-692.

Feshbach, N. D. (1969). Sex differences in children's modes of aggressive responses toward outsiders. *Merrill-Palmer Quarterly, 15*(3), 249-258.

Galen, B. R., & Underwood, M. K. (1997). A developmental investigation of social aggression among children. *Developmental Psychology, 33*(4), 589-600.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Hart, E.L., Lahey, B.B., Loeber, R., & Hanson, K.S. (1994). Criterion validity of informants in the diagnosis of disruptive behavior disorders in children: A preliminary study. *Journal of Consulting and Clinical Psychology*, 62, 410-414.

Hart, C. H., Nelson, D. A., Robinson, C. C., Olsen, S. F., & McNeilly-Choque, M. (1998). Overt and relational aggression in Russian nursery-school-age children: Parenting style and marital linkages. *Developmental Psychology, 34*(4), 687-697.

Hay, D. A., McStephen, M., Levy, F., & Pearsall-Jones, J. (2002). Recruitment and attrition in twin register studies of childhood behavior: The example of the australian twin ADHD project. *Twin Research, 5*(5), 324-328.

Henington, C., Hughes, J. N., Cavell, T. A., & Thompson, B. (1998). The role of relational aggression in identifying aggressive boys and girls. *Journal of School Psychology, 36*(4), 457-477.

Hipwell, A. E., Loeber, R., Stouthamer-Loeber, M., Keenan, K., White, H. R., & Krone-Man, L. (2002). Characteristics of girls with early onset disruptive and antisocial behaviour. *Criminal Behaviour and Mental Health, 12*(1), 99-118.

Hope, T.L., Adams, C., Reynolds, L., Powers, D., Perez, R.A., & Kelley, M.L. (1999). Parent vs. self-report: Contributions towards diagnosis of adolescent psychopathology. *Journal of Psychopathology and Behavioral Assessment*, 21, 349-363.

Iacono, W.G., Carlson, S.R., Taylor, J., Elkins, I.J., & McGue, M. (1999). Behavioral disinhibition and the development of substance use disorders: Findings from the Minnesota twin family study. *Development and Psychopathology*, 11, 869-900.

Jensen, P.S., Rubio-Stipec, M., Canino, G., Bird, H.R., Duncan, M.K., Schwab-Stone, M.E., & Lahey, B.B. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry*, 38, 1569-1579.

Johnson, D. R., & Foster, S. L. (2005). The relationship between relational aggression in kindergarten children and friendship stability, mutuality, and peer liking. *Early Education and Development, 16*(2), 141-160.

Juliano, M., Werner, R. S., & Cassidy, K. W. (2006). Early correlates of preschool aggressive behavior according to type of aggression and measurement. *Journal of Applied Developmental Psychology, 27*(5), 395-410.

Kamper, K. E., & Ostrov, J. M. (2013). Relational aggression in middle childhood predicting adolescent social-psychological adjustment: The role of friendship quality. *Journal of Clinical Child and Adolescent Psychology, 42*(6), 855-862.

Kaukiainen, A., Björkqvist, K., Lagerspetz, K., Österman, K., Salmivalli, C., Rothberg, S., & Ahlbom, A. (1999). The relationships between social intelligence, empathy, and three types of aggression. *Aggressive Behavior, 25*(2), 81-89.

Keenan, K., Coyne, C., & Lahey, B. B. (2008). Should relational aggression be included in DSM-V? *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(1), 86-93.

Klump, K.L. and S.A. Burt. (2006). The Michigan State University Twin Registry (MSUTR): Genetic, environmental, and neurobiological influences on behavior across development. *Twin Research and Human Genetics*, 9: 971-977.

Krueger, R. F., Caspi, A., & Moffitt, T. E. (2000). Epidemiological personology: The unifying role of personality in population-based research on problem behaviors. *Journal of Personality, 68*(6), 967-998.

Lagerspetz, K. M., Björkqvist, K., & Peltonen, T. (1988). Is indirect aggression typical of females? gender differences in aggressiveness in 11- to 12-year-old children. *Aggressive Behavior, 14*(6), 403-414.

McConaughy, S.H. & Achenbach, T.A. (2001). *Manual for the Semistructured Interview for Children and Adolescents*. 2nd ed., Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

McEvoy, M. A., Estrem, T. L., Rodriguez, M. C., & Olson, M. L. (2003). Assessing relational

and physical aggression among preschool children: Intermethod agreement. *Topics in Early Childhood Special Education, 23*(2), 53-63.

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In *Handbook of Structural Equation Modeling* (pp. 380-393). New York, NY: The Guilford Press.

Moretti, M. M., Holland, R., & McKay, S. (2001). Self–other representations and relational and overt aggression in adolescent girls and boys. *Behavioral Sciences & the Law, 19*(1), 109-126.

Moroschan, G., Hurd, P. L., & Nicoladis, E. (2009). Sex differences in the use of indirect aggression in adult canadians? *Evolutionary Psychology, 7*(2), 146-159.

Muthen, L. K., & Muthen, B. O. (1998-2012). *Mplus user's guide*. Seventh Edition. Los Angeles, CA: Muthen & Muthen.

Österman, K., Björkqvist, K., Lagerspetz, K. M. J., Kaukiainen, A., Landau, S. F., Frączek, A., & Caprara, G. V. (1998). Cross-cultural evidence of female indirect aggression. *Aggressive Behavior, 24*(1), 1-8.

Ostrov, J. M., & Crick, N. R. (2005). Current directions in the study of relational aggression during early childhood. *Early Education and Development, 16*(2), 109-113.

Paquette, J. A., & Underwood, M. K. (1999). Gender differences in young adolescents' experiences of peer victimization: Social and physical aggression. *Journal of Developmental Psychology, 45*(2), 242-266.

Peeters, H., Gestel, S.V., Vlietinck, R., Derom, C., & Derom, R. (1998). Validation of a telephone zygosity questionnaire in twins of known zygosity. *Behavior Genetics,* 28(3), 159-163.

Pellegrini, A. D., & Roseth, C. J. (2006). Relational aggression and relationships in preschoolers: A discussion of methods, gender differences, and function. *Journal of Applied Developmental Psychology, 27*(3), 269-276.

Prinstein, M. J., Boergers, J., & Vernberg, E. M. (2001). Overt and relational aggression in adolescents: Social–psychological adjustment of aggressors and victims. *Journal of Clinical Child Psychology, 30*(4), 479-491.

Putallaz, M., Grimes, C. L., Foster, K. J., Kupersmidt, J. B., Coie, J. D., & Dearing, K. (2007). Overt and relational aggression and victimization: Multiple perspectives within the school setting. *Journal of School Psychology, 45*(5), 523-547.

Rose, A. J., Swenson, L. P., & Carlson, W. (2004). Friendships of aggressive youth: Considering

the influences of being disliked and of being perceived as popular. *Journal of Experimental Child Psychology, 88*(1), 25-45.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132*(1), 98-131.

Salmivalli, C., & Kaukiainen, A. (2004). "Female aggression" revisited: Variable- and person-centered approaches to studying gender differences in different types of aggression. *Aggressive Behavior, 30*(2), 158-163.

Samuel, D. B., Carroll, K. M., Rounsaville, B. J., & Ball, S. A. (2013). Personality disorders as maladaptive, extreme variants of normal personality: Borderline personality disorder and neuroticism in a substance using sample. *Journal of Personality Disorders, 27*(5), 625-635.

Scheithauer, H., Haag, N., Mahlke, J., & Ittel, A. (2008). Gender and age differences in the development of relational/indirect aggression: First results of a meta-analysis. *European Journal of Developmental Science, 2*(1-2), 176-189.

Spieker, S. J., Campbell, S. B., Vandergrift, N., Pierce, K. M., Cauffman, E., Susman, E. J., & Roisman, G. I. (2012). Relational aggression in middle childhood: Predictors and adolescent outcomes. *Social Development, 21*(2), 354-375.

Stark, S., Chernyshenko, O.S., & Drasgow, F. (2004) Examining the effects of differential item (functioning and differential) test functioning on selection decisions: When are statistically significant errors practically important? *Journal of Applied Psychology, 89,* 497-508.

Suzuki, T., Samuel, D. B., Pahlen, S., & Krueger, R. F. (2015). DSM-5 alternative personality disorder model traits as maladaptive extreme variants of the five-factor model: An item-response theory analysis. *Journal of Abnormal Psychology,* doi:http://dx.doi.org/10.1037/abn0000035

Tay, L., Meade, A. W., Cao, M. (2015). An overview and practical guide to IRT measurement equivalence analysis. *Organizational Research Methods, 18*(1), 3-46. doi: 10.1177/1094428114553062

Tomada, G., & Schneider, B. H. (1997). Relational aggression, gender, and peer acceptance: Invariance across culture, stability over time, and concordance among informants. *Developmental Psychology, 33*(4), 601-609.

Underwood, M. K., Galen, B. R., & Paquette, J. A. (2001). Top ten challenges for understanding gender and aggression in children: Why can't we all just get along? *Social Development, 10*(2), 248-266.

Vaillancourt, T. (2005). Indirect aggression among humans: Social construct or evolutionary adaptation? *Developmental origins of aggression.* (pp. 158-177) Guilford Press, New York, NY.

van de Schoot, R., Lugtig, P., & Hox, J. (2012) A checklist for testing measurement invariance. *Developmetrics*, 1-7.

Xie, H., Cairns, R. B., & Cairns, B. D. (2002). The development of social aggression and physical aggression: A narrative analysis of interpersonal conflicts. *Aggressive Behavior, 28*(5), 341-355.

Xie, H., Cairns, R. B., & Cairns, B. D. (2002). The development of social aggression and physical aggression: A narrative analysis of interpersonal conflicts. *Aggressive Behavior, 28*(5), 341-355.

Zalecki, C. A., & Hinshaw, S. P. (2004). Overt and relational aggression in girls with attention deficit hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 125-137.