Emotions and Cognitions in Early Childhood Aggression: the Role of Irritability and Hostile Attribution Biases

Gretchen R. Perhamus [1] · Jamie M. Ostrov [1]

Abstract

The present study examines the role of irritability in form and function subtypes of aggression over 1 year in early childhood ($N =$ 300, $M_{age} = 44.70$ months, $SD = 4.38$ months). This study prospectively tests hostile attribution biases (HAB) as a mediator in irritability's relations with aggression, with hypothesized form-specific relations between HAB and aggression. Moderation by gender and a reversed alternative model (aggression to irritability, mediated by HAB), were also tested. Path analyses showed irritability predicted increases in all subtypes of aggression ($\beta s = 0.24$–0.34), but with moderation by gender, such that irritability significantly predicted increases in reactive relational aggression for girls only ($\beta = 0.43$). Reactive physical aggression also significantly predicted increases in irritability ($\beta = 0.15$). HAB was not associated significantly with any forms or functions of aggression, although gender differences emerged between HAB for instrumental provocations and reactive physical aggression. No significant indirect effects were found. Results highlight the importance of considering both forms and functions of aggression when investigating irritability, and point to potential gender differences in the role of irritability in relational aggression in early childhood.

Keywords Irritability · Aggression · Hostile attribution bias · Early childhood

From a developmental psychopathology perspective, it is important to understand how behaviors and emotional styles interact to predict problems and positive outcomes later in development (Cicchetti 2006). Additionally, understanding the range of typical development increases our comprehension of how these processes may go awry for those at the extreme end of the spectrum (e.g., psychopathology), and where intervention efforts may be targeted (Cicchetti 2006; Sroufe 2013). To this end, this study investigates a problematic behavior (i.e., aggression, defined as actions intended to hurt or harm another; Dodge et al. 2006) associated with poorer psychological and social outcomes (Coie and Dodge 1998), in a developmental period (i.e., early childhood) where it is especially common (Ostrov and Keating 2004) and potentially impactful. During early childhood, children are learning to navigate peer interactions, and establishing peer competence is a key

developmental task (Sroufe 2013). Aggressive behaviors can disrupt this process, setting the stage for future behavioral and emotional difficulties (Coie and Dodge 1998).

Aggression is known to have highly correlated but distinct relational and physical forms and reactive and proactive functions (e.g., Ostrov and Crick 2007), associated with differential developmental correlates and outcomes (e.g., Card and Little 2006; Murray-Close et al. 2016). Physical aggression involves the use or threat of physical force to harm another (e.g., kicking, hitting, shoving), whereas in relational aggression actual or threat of damage to a relationship (e.g., social exclusion) is the mechanism of harm (Crick and Grotpeter 1995). Both forms of aggression may serve different functions; they may be exhibited proactively to achieve instrumental goals (e.g., to get a desired toy), or reactively in response to a perceived threat or provocation (e.g., hitting someone who took a toy; Card and Little 2006). These forms and functions may be crossed to create four subtypes of aggression – reactive physical aggression, reactive relational aggression, proactive physical aggression, and proactive relational aggression (e.g., Ostrov and Crick 2007). An alternative analytic approach, using structural equation modeling (SEM) to statistically isolate each form and function of aggression, has also been

---

✉ Gretchen R. Perhamus
grperham@buffalo.edu

1 Department of Psychology, University at Buffalo, The State University of New York, 475 Park Hall, Buffalo, NY 14260, USA

proposed (Little et al. 2003). However, this approach creates orthogonal constructs that can be difficult to interpret, as in daily life no act of aggression occurs without both form and function (Underwood 2003). Therefore, this paper adopts the crossed four aggression subtypes approach.

Temperament and emotional tendencies, such as negative affectivity, have been implicated as linking and differentiating factors among these different aggression subtypes (e.g., Tackett et al. 2014). This paper tests irritability as one of these factors. Although definitions of irritability vary, and a variety of terms (e.g., anger proneness, frustration intolerance) are often used in the literature, there is some consensus that this construct represents a facet of reactive negative affectivity, coupled with higher approach and appetitive reward tendencies (Brotman et al. 2017; DeSerisy and Deveney 2019; Rothbart et al. 2001). In other words, irritable children show an increased sensitivity to and desire for reward, leading to a higher propensity to experience things as threatening or frustrating, and show an approach response to these frustrations and threats, expressed as the negative emotion of anger (Brotman et al. 2017).

An emotionally-integrated social-information processing (SIP) model provides a framework to examine how emotional tendencies exert influence on behavioral processes (Lemerise and Arsenio 2000). In this model, individuals engage in multiple cognitive steps when deciding on a behavioral response to a social stimulus, with temperament and mood influencing each of these steps (Lemerise and Arsenio 2000). Aggression is thought to result from biases in all areas of the SIP model (Crick and Dodge 1994), but much research has focused on Step 2, interpretation of cues (e.g., Godleski and Ostrov 2010; Orobio de Castro et al. 2002). Youth who are aggressive are thought to show a hostile attribution bias (HAB), or be more likely to interpret neutral or ambiguous cues as threatening (Crick and Dodge 1994). These attribution biases can be in response to relational (e.g., not getting invited to a birthday party, HAB-R) or instrumental (e.g., someone bumping into you from behind, HAB-I) provocations (Crick 1995; Crick et al. 2002; Godleski and Ostrov 2010; Nelson et al. 2008). Aggressors may show more HAB for situations related to the form of aggression they are exhibiting (e.g., Crick 1995; Crick et al. 2002), and past work in older samples has found these relations to be specific to reactive subtypes (i.e., HAB-R with reactive relational aggression, HAB-I with reactive physical aggression; Bailey and Ostrov 2008; Murray-Close et al. 2010).

We address three aims related to these constructs. First, we are the first known study to test prospective relations of irritability in all four subtypes of aggression. Second, HAB (i.e., HAB-R, HAB-I) is considered as a mediator in irritability's relations with aggression subtypes, with an eye to form-specific pathways. Finally, gender is tested as a potential moderator.

Irritability in Forms and Functions of Aggression

Irritability, as operationalized above, is distributed continuously in the population (Brotman et al. 2017). Developmentally normative irritability peaks in early childhood, and presents as a risk factor for a number of poorer developmental outcomes (Brotman et al. 2017; Camacho et al. 2019; DeSerisy and Deveney 2019). Theory and empirical work have demonstrated irritability and anger have stronger associations with reactive than proactive aggression (Brotman et al. 2017; Card and Little 2006; Hubbard et al. 2004; Stoddard et al. 2019). However, associations with relational forms of aggression remain understudied, and no studies to our knowledge have considered the role of irritability in crossed form and function aggressive subtypes. Therefore, it is currently unknown whether irritability's predominant associations with reactive, over proactive, functions extend to relational subtypes. In an early childhood sample, Ostrov et al. (2013) found anger to be associated with increases in proactive and reactive physical aggression and proactive relational aggression, but not reactive relational aggression (Ostrov et al. 2013). However, this study considered only phasic, physical displays of anger (i.e., behavioral outburst of intense anger; Copeland et al. 2015), and included emotion regulation skills as a separate variable in regression models (Ostrov et al. 2013). It therefore may not have captured the tonic, temperamental irritability construct (i.e., persistent grouchy, grumpy, or angry mood; Copeland et al. 2015). To our knowledge this is the first examination of tonic irritability's relations with both physical and relational forms of proactive and reactive aggression. Specifically, this study provides a novel prospective examination of irritability's associations with these four subtypes, hypothesizing it to be associated with increases in reactive physical and relational aggression.

HAB as a Mediator

The tendency of irritable youth to interpret neutral and ambiguous cues as hostile is well established (DeSerisy and Deveney 2019; Stoddard et al. 2019), and a key component of models of severe, chronic irritability (Brotman et al. 2017). Higher approach and reactive tendencies to threats, as well as increased reward sensitivity, potentiate individuals to have a lower threshold for experiencing threat, thereby increasing their propensity to interpret ambiguous situations as threatening (i.e., HAB; Brotman et al. 2017). Additionally, the emotionally integrated SIP model predicts those with higher intensity emotional responding rely more heavily on heuristics when interpreting social situations (Lemerise and Arsenio 2000). Therefore, temperamental irritability, associated with more frequent and more intense anger in threatening

situations, may cause children to use their biased attributions more often, reinforcing them over time.

In turn, HAB has also been consistently positively associated with aggression (e.g., Orobio de Castro et al. 2002), although findings have been less consistent in young children (e.g., Orobio de Castro et al. 2002; Schultz et al. 2018). This inconsistency is likely due to a combination of developmental (e.g., intent understanding, receptive and expressive language), and methodological (e.g., response bias) challenges in assessing hostile attributions in young children (see Schultz et al. 2018 for a review). Nevertheless, levels of HAB have been shown to be positively associated with externalizing behavior broadly (e.g., Schultz et al. 2018; Ziv 2012) and aggressive behavior specifically (e.g., Feshbach 1989; Katsurada and Sugawara 1998; Runions and Keating 2010) in early childhood. Higher levels of preschool HAB have also prospectively predicted higher levels of problem behaviors in early school years (e.g., Dodge et al. 1990; Runions and Keating 2007). Furthermore, recent work has demonstrated that preschoolers can respond to hostile attribution vignettes reliably and validly with appropriate developmental modifications (Schultz et al. 2018).

Importantly, no known studies of HAB's associations with aggression in early childhood have considered forms of HAB, or forms and functions of aggression. Theoretically and empirically at older ages, HAB is linked predominantly to reactive aggressive responding (e.g., Orobio de Castro et al. 2002), with some work suggesting form-specific relations as described above (e.g., Crick et al. 2002; Mathieson et al. 2011). Furthermore, two studies in emerging adulthood and adulthood, the only known to have considered both forms of HAB and form and function aggression subtypes, found subtype-specific effects (i.e., HAB-R predicted reactive relational aggression, HAB-I predicted reactive physical aggression; Bailey and Ostrov 2008; Murray-Close et al. 2010). Together, this suggests neglecting form and function distinctions in HAB's associations with aggression in prior early childhood studies may have further masked meaningful effects.

Two known studies have examined the role of related emotional processes in HAB's relations with aggression in early childhood, with conflicting results. Helmsen et al. (2012) found no association between emotion regulation or aggression with HAB, whereas Runions and Keating (2010) found HAB predicted higher levels of aggression at high, but not low, levels of dispositional anger. There are several potential explanations for these discrepant results. Helmsen et al. (2012) opted to dichotomize their HAB variable based on whether children made at least one hostile attribution, which may have reduced power to detect effects. Additionally, Runions and Keating (2010) used an HAB measure with uniquely-worded, and potentially developmentally appropriate, forced-choice response options (Schultz et al. 2018).

Notably, neither of these studies differentiated between subtypes of aggression or HAB, and both were cross-sectional and used slightly older samples than the current study. Although Runions and Keating (2010) found support for a moderation model, they found a synergistic, rather than disordinal (i.e., cross-over) effect, which does not rule out the applicability of a mediation relation, and the cross-sectional study design prevented them from testing mediation. Finally, irritability and SIP theory support a mediation relation (as described above). Therefore, this study tests a model in which HAB mediates irritability's relations with aggression. Specifically, it tests the hypothesis that irritability is positively associated with HAB, HAB in turn produces increases in reactively aggressive responses, and this process is differentiated by HAB to relational and instrumental threats (Fig. 1a).

A reversed alternative model is also tested, such that being more aggressive leads to higher HAB, in turn promoting increases in irritability (Fig. 1b). Being aggressive may provoke more frequent angry responses from others, increasing HAB over time. This is also consistent with Lemerise and Arsenio's (2000) expanded SIP model, and has been partially supported in past work (Godleski and Ostrov 2010). In turn, HAB may maintain or increase irritability through an increased propensity to experience situations as threatening (Brotman et al. 2017; Stoddard et al. 2019), a process that may be especially detectable in early childhood given changes in irritability levels from year to year (Camacho et al. 2019).

Gender Moderation

Finally, this study examines gender as a moderator of these models. Past work has suggested the modal form of aggression differs by gender, such that girls engage more often in relational than physical aggression, and boys demonstrate the reverse (e.g., Crick et al. 1996; Bailey and Ostrov 2008; Ostrov and Crick 2007), but relations with HAB have been mixed. Several studies have found no gender differences in relations between HAB and aggression (e.g., Crick et al. 2002; Runions and Keating 2010). In a middle-childhood sample, girls showed increases in HAB-R from 3rd to 6th grade, whereas boys showed increases in HAB-I (Godleski and Ostrov 2010). Further, whereas some studies have found support for gender-modal pathways (i.e., HAB-I associated with physical aggression for boys, HAB-R with relational aggression for girls; Mathieson et al. 2011; Nelson et al. 2008), other work has suggested those exhibiting the non-modal form of aggression show the highest levels of HAB (e.g., Godleski and Ostrov 2010). Due to these inconsistent prior findings and that this is the first test of gender differences with HAB as a mediator of irritability and aggression, this aim is exploratory.
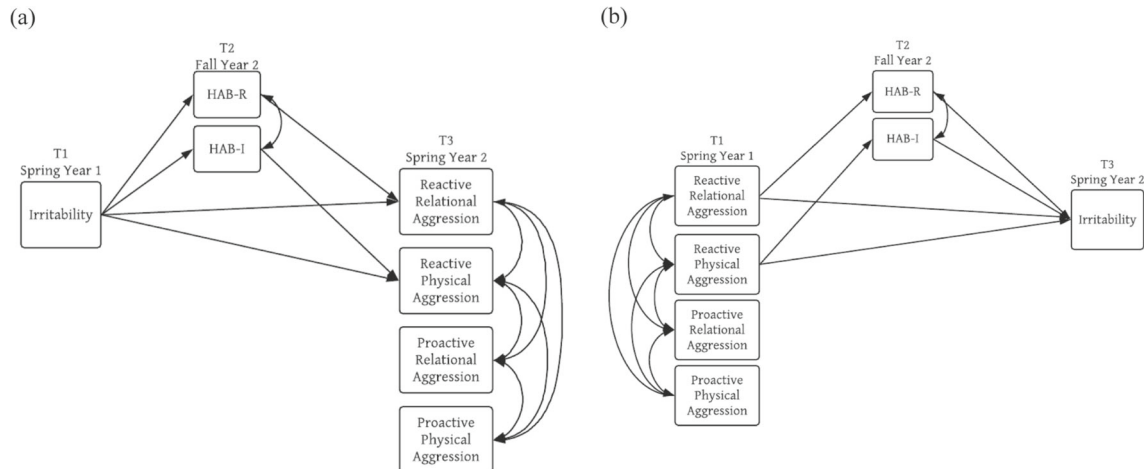
Hypothesized Conceptual Models

Fig. 1 Hypothesized Conceptual Models. *Note.* Primary model is depicted in (**a**), Alternative model in (**b**). HAB-R = Hostile attribution bias for relational provocations, HAB-I = Hostile attribution bias for instrumental provocations. T1 levels of outcome variables, age, and school SES included as covariates. Only hypothesized paths are shown

Current Study

In sum, this study tests three complimentary aims. First, it tests irritability's relations with four subtypes of aggression, taking into account both form and function, hypothesizing that irritability will be associated with increases in reactive physical and reactive relational aggression. Second, it tests forms of HAB as a mediating social-cognitive process in these associations, predicting HAB-R to be associated with reactive relational aggression and HAB-I with reactive physical aggression. Finally, it tests potential moderating effects of gender. To our knowledge, this study provides the first prospective test of irritability's associations with aggression considering both form and function. Likewise, this is the first test of HAB's mediational role in this or related (i.e., dispositional anger, emotion regulation) constructs taking into account forms and functions of aggression and forms of HAB. This is done at an especially relevant developmental time (i.e., early childhood), when both irritability and aggression are prevalent, easily observed, and potentially developmentally impactful (DeSerisy and Deveney 2019; Ostrov and Keating 2004). Finally, this study uses a prospective, multi-method design, particularly well-suited for testing mediation models.

Methods

Participants and Procedure

All study procedures were approved by the local institutional review board (IRB) of University at Buffalo, State University of New York. Four cohorts were recruited over a four-year period through partnerships with ten National Association for the Education of Young Children (NAEYC) accredited or recently accredited early childhood education centers in a large northeastern city and surrounding suburbs. Each year, all children in participating preschool (i.e., 3–5 year old) classrooms were invited to join the study through consent forms distributed to families by teachers. Approximately 56% of eligible families returned consent forms to participate in the study (see Ostrov et al. 2019 for additional details). Cohorts were merged to create the final sample consisting of 300 preschoolers (M_{age} at T1 = 44.70 months, SD = 4.38 months, n = 132 female). The sample is middle to upper-middle class on average, and represents relatively diverse ethnic and racial backgrounds (3.0% African American/Black, 7.6% Asian/Asian American/Pacific Islander, 1.0% Hispanic/Latinx, 11.3% multi-racial, 62.1% White, and 15.0% missing/unknown).

Data was collected from each cohort at three time points over two academic years [spring/summer of year 1 (T1), fall of year 2 (T2), and spring of year 2 (T3)]. T1 included school-based observations, teacher and parent questionnaires, and a lab-based child interview. T2 included teacher reports and a school-based child interview, and T3 included teacher reports and school-based observations. Children provided verbal assent prior to all interviews, and teachers provided written consent before completing teacher reports.

Measures

Irritability

Irritability was constructed as a composite from teacher reports on two questionnaires assessing anger and frustration. The Anger/Frustration subscale of the Child Behavior

Questionnaire – Short Form (CBQ-TF; Putnam and Rothbart 2006; Rothbart et al. 2001) contains six items assessing tendency to experience anger and frustration scored on a 1 (extremely untrue) to 7 (extremely true) scale, averaged to create a subscale score. The scale contains items assessing tendency to experience frustration generally (e.g., "Gets angry when s/he can't find something s/he wants to play with") and within the context of limit-setting (e.g., "Gets quite frustrated when prevented from doing something s/he wants to do"). Both types of items can be considered representations of this study's operationalization of irritability, as they depict an angry approach response to either blocked or withheld reward. For example, frustration when not finding something to play with can be considered an angry response to lack of reward (i.e., having a fun time playing) when expected (i.e., during play time). The subscale has been previously validated (Putnam and Rothbart 2006), and has been used in studies measuring similar constructs in this age group (e.g., Camacho et al. 2019; Runions and Keating 2010).

Four items assessing displays of anger, adapted from an observational method by Hubbard et al. (2004), were also incorporated. The items include "expresses anger with peers," "gets angry during play," "uses toys or classroom materials roughly (e.g., throwing toys or slamming toys down when frustrated)," and "displays frustration (e.g., swinging fist, hitting objects, hitting one's own head with the palm of the hand)". Each item is responded to on a 1 (never) to 4 (almost always) scale and are averaged to create a subscale score.

Both measures showed good internal consistency at T1 and T3 (αs = 0.82–0.91), and moderate stability across time points (rs = 0.46–0.49, ps < 0.001). Additionally, teacher reports on both forms were significantly correlated with parent report (rs = 0.17–0.24, ps < 0.05). The scales were moderately correlated with each other at T1 (r = 0.67, p < 0.001) and T3 (r = 0.67, p < 0.001). Subscale scores on both forms were standardized, and averaged to create irritability composites.

Hostile Attribution Biases

Hostile attribution biases (HAB) were assessed through child interviews by trained graduate students using a modified measure of the Assessment of Intent Attributions (Crick 1995), based on child hostile attribution bias assessments (Casas and Crick 1999; Crick 1995), further adapted for young children (Godleski and Ostrov 2020). The interview includes eight brief ambiguous vignettes, including four relational (e.g., not receiving an invitation to a birthday party) and four instrumental (e.g., being bumped from behind while standing in line) provocation situations. Several steps were taken to support the developmental appropriateness of the task. First, to simplify the response task, children were asked to initially indicate whether they think the person in the story was "trying to be mean" or "not trying to be mean". If they indicate "trying to be mean," the child is then asked if they were being "really mean" or "a little mean". Second, the procedure called for the order of response options to be randomized across vignettes to mitigate the documented tendency of young children to disproportionately select the last presented option (Schultz et al. 2018). Third, each vignette was intentionally brief, began with "Let's pretend," and focused on familiar contexts (e.g., birthday party, playground). Finally, all responses were made using a picture board with faces representing each response option, and interviewers verbally confirmed each response. Responses are scored from 0 (not trying to be mean) to 2 (really mean), and summed for each subscale. The present measure showed acceptable internal consistency for both HAB subtypes with one story removed (αs = 0.69–0.72) and showed convergent validity between parent and child levels of HAB in an early childhood sample (Godleski and Ostrov 2020). In the current study, internal consistency was lower than convention but acceptable for HAB-I (α = 0.62) and HAB-R (α = 0.63) with the first item of each subscale removed as in previous work (Godleski and Ostrov 2020), and was consistent with prior work using similar measures with young children (e.g., Runions and Keating 2010).

Forms and Functions of Aggression

A composite of naturalistic observations and observer report was used to assess forms and functions of aggression. Observations were completed in a school setting by trained undergraduate and graduate-level research assistants at T1 and T3. Using a focal child sampling with continuous recording procedure (Ostrov and Keating 2004; Crick et al. 2006), each child was observed during free play for eight ten-minute sessions over an eight-week period (80 min total at each time point). Before conducting observations, research assistants spent time participating in classrooms to minimize reactivity. Observers recorded instances of children engaging in relational and physical aggression. These observations were then coded as instances of proactive or reactive relational or physical aggression as four exclusive categories. Context and sequences of interactions were taken into account as necessary (i.e., presence of a desired goal for proactive aggression, behavior occurring after a perceived threat for reactive aggression). The number of times children engaged in these behaviors across observation sessions was averaged within behavior type to create an observation aggression score.

Past work with this technique has established evidence for both inter-rater reliability and validity (Ostrov and Crick 2007; Ostrov and Keating 2004). In the current study, 15% of classroom observations and 50% of proactive/reactive secondary codes were independently coded by a second observer and assessed for reliability. Direct observations were reliable across reporters (ICCs > 0.70). Secondary codes of aggression

functions showed acceptable reliability (Cohen's κs = 0.60–0.88; Pellegrini 2004), with the exception of T1 proactive relational aggression (κ = 0.50), which deserves caution. However, given the stringent nature of κ (Pellegrini 2004), that these levels are similar to those in past work using this coding method (Ostrov and Crick 2007), and that observational methods help distinguish independent effects between aggression subtypes (an important goal of this study; Card and Little 2006), these observations are retained.

Following the conclusion of observations at each time point, one randomly selected observer from each classroom completed a psychometrically strong teacher report measure of forms and functions of aggression (Preschool Proactive and Reactive Aggression-Teacher Report, PPRA-TR-R; Ostrov and Crick 2007). This measure includes three items assessing each subtype of aggressive behavior (e.g., Proactive physical - "This child often hits, kicks or pushes to get what s/he wants"; Reactive relational – "When this child is upset with others, s/he will often ignore or stop talking to them"), rated from 1 (Never or Almost Never True) to 5 (Always or Almost Always True). Items were averaged within aggression subtype. RA's completed these ratings after spending several months observing children's behavior, and past work has found RA's to be reliable and valid reporters using similar procedures (e.g., Murray-Close and Ostrov 2009). The ratings showed good internal consistency (αs = 0.79–0.91) in the current study.

Observations and RA reports were standardized and averaged within aggression subtype to create composites. This addressed concerns regarding restricted range in observations and high intercorrelation of subtypes on the observer reports (rs = 0.83–0.86, ps < 0.001). RA reports and observations were significantly correlated within aggression subtype at both time points (rs = 0.17–0.32, ps < 0.05). Due to administrative error, RA report was not obtained for one cohort at T1 (n = 126) and another at T3 (n = 18). Standardized observation scores only were used for these participants at these time points.

Analytic Plan

Data were subject to a number of cleaning procedures prior to data analysis. First, descriptive statistics of measures were obtained, and an analysis of outliers and potential non-normality was conducted. Outliers were defined as scores further than 3 standard deviations from the mean and were adjusted to the outer bound of this limit (Kline 2016). Skew ranged from 0.15 to 2.42 and kurtosis from −0.94 to 5.88, indicating no non-normality concerns (Kline 2016).

Data were also assessed for systematic missingness. Data was missing for 31.7% (n = 95) of participants from T1 to T2, and 30.3% (n = 91) from T1 to T3. Most attrition occurred during the transition between school years, with low attrition

from T2 to T3 (1.3%). This was anticipated given the nature of the study, and was primarily due to children moving to kindergarten, changing schools for reduced cost universal pre-kindergarten programs, or moving from the area. Little's (1988) MCAR test suggested the data were not missing completely at random [$\chi^2(186) = 245.60$, $p = 0.002$]. Missingness was not associated with significant differences on key predictor or outcome variables, gender, or age. Lower SES was associated with missingness at T2 [$t(82.70) = -2.27$, $p = 0.03$; $d = 0.37$] and T3 [$t(68.82) = -2.89, p = 0.005; d = 0.48$]. However, SES was assessed using Hollingshead codes of parent occupation, a limited measure, and was unavailable for 24% of participants (n = 73). School code was also associated with missingness at T2 [$\chi^2(9) = 26.84$, $p = 0.001$; Cramer's $V = 0.30$] and T3 [$\chi^2(9) = 38.54$, $p < 0.001$; Cramer's $V = 0.36$]. Further, schools differed significantly in proportion of high vs. low SES participants based on a median split in occupation code [$\chi^2(9) = 20.99, p = 0.01$; Cramer's $V = 0.27$], and school code was available for all participants. Therefore, schools were rank ordered by proportion of high vs. low SES participants, and this rank-ordered school code was controlled in subsequent models as a proxy for SES. As missingness was related to a known variable, missing data was accommodated using full information maximum likelihood (FIML; Little 2013).

Correlations and variance inflation factor (VIF) results were analyzed for multicollinearity concerns across HAB and child aggression subtypes. Values did not suggest a need to collapse across subtypes (VIF values = 1.19–3.10; rs < 0.13–0.54), with the exception of the correlation between proactive and reactive physical aggression at T3 ($r = 0.72$, $p < 0.001$). However, given the centrality of including form and function of aggression in this study, that no other forms or functions of aggression at T1 or T3 were highly corelated, and that this value was only just above a cutoff of 0.70, we opted not to collapse across functions of aggression.

Path analyses were conducted in Mplus Version 8.4 (Muthén and Muthén 1998–2020) with maximum likelihood estimation (ML), using indirect effects testing to test mediation models. For the primary model (Fig. 1a), proactive and reactive physical aggression were regressed onto HAB-I, and proactive and reactive relational aggression were regressed onto HAB-R, which were both then regressed onto T1 irritability. For the alternative model (Fig. 1b), T3 irritability was regressed onto HAB-I and HAB-R, which were regressed onto physical and relational aggression subtypes, respectively. Next, gender was entered as a grouping variable, and models were run with paths constrained to equivalence across gender. Paths were then freed, and improvement in model fit for the unconstrained model was assessed using change in model χ^2. If the χ^2 difference test indicated significant improvement in model fit, regression paths were examined for differences across gender using modification indices (MI) to determine

which parameters should be sequentially freed in accordance with procedures outlined by Yoon and Millsap (2007). Age at T1, school, and initial levels of outcome variables were entered as covariates at all points of models. HAB and aggression subtypes were allowed to covary.

Overall model fit was assessed using a likelihood ratio χ^2 test with $p > 0.05$ indicating good model fit. The comparative fit index (CFI; Bentler 1990), of which values > 0.95 suggest good fit, the standardized root mean square residual (SRMR), where values < 0.05 represent good fit (Hu and Bentler 1999), and the root mean square error of approximation (RMSEA; Steiger 1990), where values < 0.05 indicate close fit (Browne and Cudeck 1992; MacCallum et al. 1996) also were considered. A total of 5000 bootstrap samples and 95% bias-corrected confidence intervals were used to test indirect effects (Hayes and Preacher 2010).

Results

Preliminary Analyses

Descriptive information and bivariate correlations are presented in Table 1. Irritability and proactive and reactive physical aggression showed moderate stability from T1 to T3. At the bivariate level, both T1 and T3 levels of irritability were associated significantly and positively with both functions of physical aggression at both time points, and both functions of relational aggression at T3. HAB-I was positively associated with irritability at T3, whereas HAB-R was not associated with irritability at either time point. Neither form of HAB was associated with any aggression subtypes at either time point. Age at T1 was positively associated with reactive relational aggression at T1, proactive relational aggression at T3, and HAB-R at T2. Gender was associated significantly with physical aggression and irritability at T1 and T3. Boys showed higher levels of proactive physical aggression at T1 [t (294) = 2.10, $p = 0.04$; $d = 0.25$] and T3 [t (205) = 2.92, $p = 0.004$; $d = 0.42$] and reactive physical aggression at T1 [t (294) = 2.64, $p = 0.01$; $d = 0.31$] and T3 [t (205) = 3.40, $p = 0.001$; $d = 0.48$]. Finally, boys showed higher levels of irritability at both T1 [t (291) = 2.38, $p = 0.02$; $d = 0.28$] and T3 [t (206) = 4.75, $p < 0.001$; $d = 0.67$].

Primary Model

The primary mediation model showed close fit to the data [$\chi^2(4) = 3.85$, $p = 0.43$, CFI = 1.00, RMSEA = 0.00, SRMR = 0.01]. Irritability at T1 predicted increases in all subtypes of aggression. There was a non-significant trend toward irritability predicting higher HAB-I ($\beta = 0.17$, 95% CI [−0.02, 0.29], $p = 0.07$). There were no direct effects of irritability to HAB-R ($\beta = 0.11$, 95% CI [−0.05, 0.27], $p = 0.17$),

HAB-I to either function of physical aggression (βs = −0.01, 0.05, ps = 0.49–0.87), or HAB-R to either function of relational aggression (βs = −0.11, −0.01, ps = 0.10–0.88). Likewise, confidence intervals for indirect effects all included 0, indicating no evidence for mediation by HAB-R or HAB-I to any forms or functions of aggression.

Next, gender was entered into the model as a grouping variable, with paths of interest constrained to equivalence. Fit statistics generally suggested the constrained model provided an adequate fit to the data [$\chi^2(18) = 27.62$, $p = 0.07$, CFI = 0.98, RMSEA = 0.06, SRMR = 0.03]. Freeing the constrained paths significantly improved model fit [$\Delta\chi^2(10) = 18.53$, $p = 0.047$], suggesting paths were moderated significantly by gender. An examination of modification indices (MI) revealed two paths with MI values greater than 3.84, indicating there would be significant ($p < 0.05$) reduction in the χ^2 value if those parameters were freed (Whittaker 2012). The parameter from T1 irritability to T3 reactive relational aggression had the highest theoretically meaningful MI (MI = 8.82). After freeing this path, the MI was also significant for freeing the path from T2 HAB-I to T3 reactive physical aggression (MI = 4.53). With both paths freed, the model provided excellent fit to the data [$\chi^2(16) = 12.99$, $p = 0.67$, CFI = 1.00, RMSEA = 0.00, SRMR = 0.03], and there was no difference in fit between this model and the fully freed model [$\Delta\chi^2(8) = 5.06$, $p = 0.75$]. T1 irritability predicted increases in reactive relational aggression for girls ($\beta = 0.43$, 95% CI [0.22, 0.60], $p < 0.001$), but not boys ($\beta = 0.05$, 95% CI [−0.13, 0.22], $p = 0.56$). Effects of HAB-I on reactive physical aggression were non-significant, but in opposite directions, for boys ($\beta = 0.12$, 95% CI [−0.04, 0.27], $p = 0.13$) and girls ($\beta = −0.10$, 95% CI [−0.27, 0.08], $p = 0.23$). Therefore, although the significant MI indicates these slopes are significantly different from each other, neither is significantly different from zero, and all indirect effects remained non-significant. Standardized coefficients of the finalized model are shown in Fig. 2a.

Alternative Model

Results from the alternative model are presented in Fig. 2b. The model provided a good fit to the data [$\chi^2(4) = 2.46$, $p = 0.65$, CFI = 1.00, RMSEA = 0.00, SRMR = 0.01]. T1 levels of reactive physical aggression significantly predicted increases in irritability ($\beta = 0.15$, 95% CI [0.02, 0.27], $p = 0.02$). Non-significant trends emerged toward reactive relational aggression predicting decreases in irritability ($\beta = −0.10$, 95% CI [−0.22, 0.001], $p = 0.06$), and proactive physical aggression ($\beta = 0.10$, 95% CI [−0.01, 0.20], $p = 0.07$), and HAB-I ($\beta = 0.12$, 95% CI [0.00, 0.23], $p = 0.06$) predicting increases in irritability. HAB-R was not associated with change in irritability ($\beta = −0.01$, 95% CI [−0.14, 0.15],

Table 1 Bivariate correlations and descriptive statistics of key variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. School	—													
2. Age (months)	−0.08	—												
3. T1 Irr	−0.10	−0.03	—											
4. T3 Irr	−0.05	−0.01	0.54***	—										
5. T2 HAB-I	0.05	0.00	0.09	0.16*	—									
6. T2 HAB-R	0.08	0.17*	0.08	0.09	0.42***	—								
7. T1 Pro PAgg	0.20**	−0.09	0.17**	0.17*	0.02	0.04	—							
8. T1 Rea PAgg	0.26**	−0.11	0.25***	0.27***	−0.04	0.01	0.48***	—						
9. T1 Pro RAgg	0.09	0.01	0.08	0.02	0.03	0.08	0.17**	0.13*	—					
10. T1 Rea RAgg	0.13*	0.13*	0.12*	−0.02	−0.06	0.06	0.19**	0.13*	0.17**	—				
11. T3 Pro PAgg	0.12	0.04	0.40***	0.34***	0.07	−0.01	0.32***	0.38***	0.11	0.12	—			
12. T3 Rea PAgg	0.17*	0.10	0.36***	0.36***	0.10	0.05	0.25***	0.32***	0.11	0.19**	0.72***	—		
13. T3 Pro RAgg	0.12	0.21*	0.23**	0.19**	0.04	−0.05	0.09	0.13	0.09	0.19**	0.43***	0.43***	—	
14. T3 Rea RAgg	0.08	0.01	0.26**	0.16*	0.09	−0.01	0.08	0.08	0.04	0.09	0.34***	0.28***	0.43***	—
M	—	44.70	0.00	0.00	2.83	2.75	0.06	0.04	0.03	0.04	0.01	0.01	0.03	−0.01
SD	—	4.38	0.91	0.91	1.79	1.80	0.94	0.96	0.95	0.98	0.83	0.82	0.84	0.79
Range	—	33.83–61.97	−1.28 - 2.66	−1.36 - 2.66	0.00–6.00	0.00–6.00	−0.82 - 3.48	−0.59 - 3.34	−0.72 - 3.91	−0.91 - 3.81	−0.84 - 3.25	−0.61 - 3.45	−0.69 - 3.47	−0.75 - 2.85

Note. T1 = Time 1, T2 = Time 2, T3 = Time 3, Irr = Irritability, HAB-I = Hostile attribution bias for instrumental provocations, HAB-R = Hostile attribution bias for Relational provocations, Pro = Proactive, Rea = Reactive, PAgg = physical aggression, RAgg = relational aggression. $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$

$p = 0.95$), no subtype of aggression was associated with either form of HAB (βs = -0.09–0.05, ps = 0.28–0.76), and no significant indirect effects emerged. With gender entered as a grouping variable, the constrained model showed adequate fit to the data [$\chi^2(18) = 26.68$, $p = 0.09$, CFI = 0.93, RMSEA = 0.06, SRMR = 0.04], and allowing paths to vary across gender did not significantly improve model fit [$\Delta\chi^2(10) = 12.62$, $p = 0.25$], indicating no gender moderation.

Discussion

This study tested three complimentary aims related to prospective relations of irritability, subtypes of aggression, and HAB in early childhood using two prospective models assessing change over one calendar year. Irritability predicted increases in all subtypes of aggression, with significant moderation by gender, such that irritability predicted increases in reactive relational aggression for girls only. Conversely, reactive physical aggression predicted increases in irritability. However, neither form of HAB was associated significantly with irritability or any subtype of aggression, and did not mediate irritability's associations with aggression.

First, this study provided the first known examination of prospective associations of irritability with crossed form and function subtypes of aggression. Based on theory and prior work, irritability was expected to be especially associated with increases in reactive functions of aggression (e.g., Hubbard et al. 2004; Stoddard et al. 2019). However, past work in early childhood also failed to find hypothesized function-specific relations between anger and aggression subtypes (Ostrov et al. 2013). Irritability and related constructs may therefore represent a general risk factor for aggressive behavior in early childhood, with differences between functions becoming evident at later ages (Song et al. 2020). Notably, only reactive physical aggression predicted increases in irritability, suggesting this subtype may play an important role in predicting higher levels of irritability over development.

Importantly, irritability predicted increases in reactive relational aggression for girls only. Relational aggression is the modal form of aggression for girls (e.g., Crick et al. 1996; Ostrov and Crick 2007). Girls tend to find relational conflict situations more distressing than boys (Crick 1995; Crick et al. 2002), and therefore may be more likely than boys to use relational aggression reactively when emotionally aroused. These findings highlight the critical importance of including relational aggression when investigating the role of irritability in aggressive behavior to capture properly the experiences of girls, which has been neglected in the irritability literature.

Hypotheses surrounding mediation of irritability's associations with aggression by HAB were not supported. Despite significant bivariate associations between HAB-I and irritability, no significant associations emerged in mediation models with either subtype of HAB. This was unexpected, and is inconsistent with SIP theory (Crick and Dodge 1994; Lemerise and Arsenio 2000), and past work in older samples (e.g., Bailey and Ostrov 2008; Crick et al. 2002; Orobio de Castro et al. 2002). However, associations of HAB with aggression have been inconsistent in early childhood and HAB is challenging to measure during this developmental period (e.g., Schultz et al. 2018). The present study adds to this equivocal literature and we call for more research using developmentally appropriate measures. Of note, although paths did not reach significance for either gender, there were significant gender differences in associations of HAB-I with

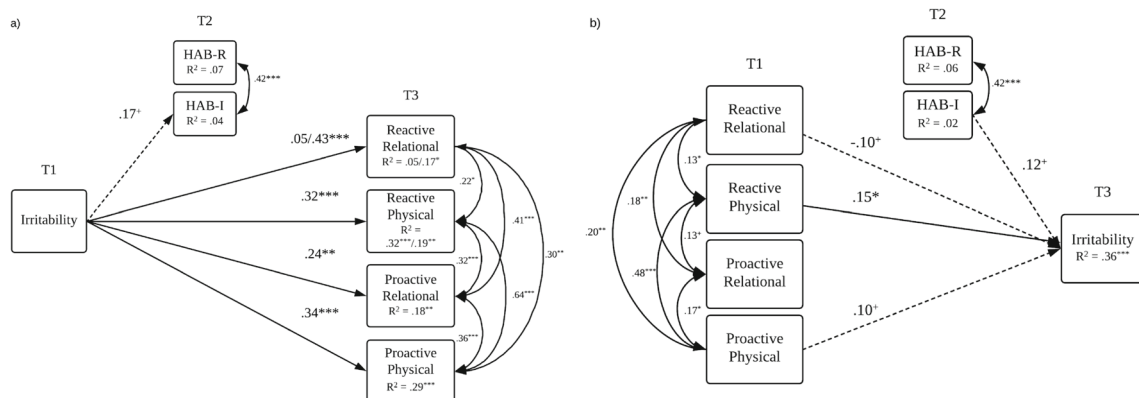*Standardized regression coefficients for mediation models.*

Fig. 2 Standardized regression coefficients for mediation models. *Note.* Primary model is depicted in (**a**), Alternative model in (**b**). HAB-R = Hostile attribution bias for relational provocations, HAB-I = Hostile attribution bias for instrumental provocations, T1 = Time 1, T2 = Time 2, T3 = Time 3. Dashed lines represent nonsignificant trends. Solid lines are significant. Nonsignificant paths are not shown. Slash values represent moderation of paths by gender, shown boys/girls. T1 levels of outcome variables, age, and school SES statistically controlled. $^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

reactive physical aggression. These results deserve continued investigation and replication.

Limitations

There are a number of limitations that should be considered in interpreting these results. First, our measurement of HAB had internal consistency values that were lower than convention, potentially limiting our ability to detect effects of HAB in the current study. Although previous work with the scale showed adequate reliability (Godleski and Ostrov 2020), in both that study and the current study, the first item in each subscale had to be dropped to achieve reliability approximating conventionally adequate levels. Additionally, despite efforts to maximize the developmental-appropriateness of the measure (e.g., counter-balancing presentation of response options, use of a picture board to illustrate levels of "meanness"), additional adaptations could be needed. For example, adding an intent understanding question, changing the wording of forced-choice response options, or using open-ended questions may help improve both the reliability and predictive validity of this measure in early childhood (Schultz et al. 2018).

Additionally, due to an administrative error, observer reports of aggression were not collected at T1 for one cohort and T3 for another, smaller cohort. For these individuals, standardized observations were used rather than a composite of observations and observer report. This may have introduced error in estimates of associations with aggression, especially at T1. Furthermore, although codes of proactive and reactive functions of aggression were generally within a reliable range, proactive relational aggression at T1 was below conventional levels. Inconsistences in this observation data across reporters may have introduced further error to our measurement of aggression. Finally, correlations between observer reports and standardized observations were relatively low, which may further limit reliability of composites.

These concerns may be compounded by our sample size, which was limited given the model complexity, especially for gender-specific models. Uneven attrition led to more boys than girls with data at T3 ($n = 91$ female, $n = 116$ male). Although missing data was accommodated using FIML, we were likely underpowered to detect small to medium effects in these models, especially for girls. In addition, due to the limited numbers of children within classrooms (e.g., 40% of T2 classrooms had 1–2 participants), we were unable to account for potential classroom nesting effects. Finally, to decrease model complexity given the sample size, forms of HAB were not tested with alternative forms of aggression (e.g., HAB-R with physical aggression), and therefore conclusions cannot be drawn surrounding these associations.

The present sample was also relatively homogeneous in SES and race/ethnicity, were recruited from high-quality child care centers, and represent a modest subset of eligible families who responded to recruitment efforts (approximately 56%), which may limit generalizability. Likewise, levels of irritability and aggression were likely primarily developmentally normative, and findings may not generalize to clinically impaired samples.

Implications and Future Directions

Despite these limitations, this study has a number of novel findings with implications for future work. The use of a longitudinal design with independent reporters for each construct of interest provided a rigorous test of change in aggression and irritability. Further, including all forms and functions of aggression and forms of HAB, as well as correlations between subtypes, allowed for a unique, stringent test of specific associations between constructs of interest.

To our knowledge, this is the first study to examine irritability's associations with aggression accounting for crossed forms and functions in early childhood. This is critically important given theorized differential associations between irritability and functions of aggression (e.g., Stoddard et al. 2019) and the critical nature of the early childhood period for learning to navigate social relationships (e.g., Sroufe 2013). Moderate stability of irritability over one calendar year indicates it can be reliably measured at this age, while also highlighting the presence of measurable change in irritability in early childhood (Camacho et al. 2019). Irritability predicted increases in all subtypes of aggression, highlighting its significance in the development of aggressive behavior. Reactive physical aggression also predicted increases in irritability, and future work should continue to investigate reciprocal relations between irritability and aggression. That irritability predicted increases in reactive relational aggression for girls specifically underlines the importance of this subtype to capture girls' experiences, which has been largely neglected in the irritability and related clinical literature (Sukhodolsky et al. 2016).

The present study's lack of associations with HAB in mediation models contributes to the equivocal early childhood HAB literature. Despite previous findings of expected associations in this developmental period (e.g., Runions and Keating 2010), and evidence of preschoolers' ability to report reliably and validly on HAB (e.g., Schultz et al. 2018), developmental limitations in intent understanding and language development (i.e., expressive and receptive language abilities may impact vignette comprehension and response) make this a particularly difficult construct to assess in early childhood (Schultz et al. 2018). Notably, this study was the second to present findings using a novel interview for assessing HAB with preschoolers (Godleski and Ostrov 2020), which has shown some promising evidence of validity (i.e., convergent validity between parent and child HAB in Godleski and Ostrov 2020; bivariate associations with irritability in the present study) and reliability. This measure should continue to be refined. Future work should continue examining other SIP steps, such

as generation and positive evaluation of aggressive responses, which have been implicated in early childhood problem behaviors (e.g., Ziv 2012), and may be more easily reported on by young children due to their more concrete nature.

Conclusion

Given long-term negative impacts of both irritability and aggression on a variety of psychosocial outcomes across development (e.g., DeSerisy and Deveney 2019; Murray-Close et al. 2016), understanding mechanisms involved in their development and maintenance has important implications for intervention and prevention efforts. This study provided a novel test of irritability's short-term longitudinal associations with aggression in early childhood, considering both form and function of aggression. Results also point to potential gender differences in the role of irritability in the development of aggression in early childhood.

Compliance with Ethical Standards

Conflict of Interest/Competing Interests None.

Consent to Participate Written informed consent was obtained from the parents and teachers of participants and verbal assent was gathered from children where appropriate.

Ethics Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the university Institutional Review Board.

Code Availability Not applicable.

References

Bailey, C. A., & Ostrov, J. M. (2008). Differentiating forms and functions of aggression in emerging adulthood: Associations with hostile attribution biases and normative beliefs. *Journal of Youth and Adolescence, 37*, 713–722. https://doi.org/10.1007/s10964-007-9211-5.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Brotman, M. A., Kircanski, K., Stringaris, A., Pine, D. S., & Leibenluft, E. (2017). Irritability in youths: A translational model. *American Journal of Psychiatry, 174*, 520–532. https://doi.org/10.1176/appi.ajp.2017.16070839.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park: Sage.

Camacho, M. C., Wakschlag, L. S., & Perlman, S. B. (2019). Early childhood irritability: Using a neurodevelopmental framework to inform clinical understanding. In A. K. Roy, M. A. Brotman, & E. Leibenluft (Eds.), *Irritability in pediatric psychopathology* (pp. 73–93). New York: Oxford University Press.

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development, 30*, 466–480. https://doi.org/10.1177/0165025406071904.

Casas, J. F., & Crick, N. R. (1999). *Assessment of preschool social information processing (APSIP)*. Unpublished manual. University of Minnesota.

Cicchetti, D. (2006). Development and psychopathology. In D. Cicchetti & D. Cohen (Eds.), *Developmental psychopathology: Theory and method* (Vol. 1, 2nd ed., pp. 1–23). New York: Wiley.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 779–862). Hoboken: John Wiley & Sons Inc..

Copeland, W. E., Brotman, M. A., & Costello, E. J. (2015). Normative irritability in youth: Developmental findings from the great smoky mountains study. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*, 635–642. https://doi.org/10.1016/j.jaac.2015.05.008.

Crick, N. R. (1995). Relational aggression: The role of intent attributions, feelings of distress, and provocation type. *Development and Psychopathology, 7*, 313–322. https://doi.org/10.1017/S0954579400006520.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74–101. https://doi.org/10.1037/0033-2909.115.1.74.

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66*, 710–722. https://doi.org/10.2307/1131945.

Crick, N. R., Bigbee, M. A., & Howes, C. (1996). Gender differences in children's normative beliefs about aggression: How do I hurt thee? Let me count the ways. *Child Development, 67*, 1003–1014. https://doi.org/10.1111/j.1467-8624.1996.tb01779.x.

Crick, N. R., Grotpeter, J. K., & Bigbee, M. A. (2002). Relationally and physically aggressive children's intent attributions and feelings of distress for relational and instrumental peer provocations. *Child Development, 73*, 1134–1142. https://doi.org/10.1111/1467-8624.00462.

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*, 254–268. https://doi.org/10.1016/j.appdev.2006.02.006.

DeSerisy, M., & Deveney, C. M. (2019). Behavioral and psychophysiological investigations of irritability. In A. K. Roy, M. A. Brotman, &

E. Leibenluft (Eds.), *Irritability in pediatric psychopathology* (pp. 109–214). New York: Oxford University Press.

Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). Mechanisms in the cycle of violence. *Science, 250*, 1678–1683. https://doi.org/10.2307/2878540.

Dodge, K.A., Coie, J. D., & Lynam, D. (2006). Aggression and antisocial behavior in youth. In W. Damon (series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., pp. 719-788). Hoboken: Wiley.

Feshbach, L. E. (1989). Aggression-conduct problems, attention-deficits hyperactivity, play, and social cognition in four-year-old boys. Unpublished doctoral dissertation, University of Washington, Seattle.

Godleski, S. A., & Ostrov, J. M. (2010). Relational aggression and hostile attribution biases: Testing multiple statistical methods and models. *Journal of Abnormal Child Psychology, 38*, 447–458. https://doi.org/10.1007/s10802-010-9391-4.

Godleski, S. A., & Ostrov, J. M. (2020). Parental influences on child report of relational attribution biases during early childhood. *Journal of Experimental Child Psychology, 192*, 104775. https://doi.org/10.1016/j.jecp.2019.104775.

Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research, 45*, 627–660. https://doi.org/10.1080/00273171.2010.498290.

Helmsen, J., Koglin, U., & Petermann, F. (2012). Emotion regulation and aggressive behavior in preschoolers: The mediating role of social information processing. *Child Psychiatry & Human Development, 43*, 87–101. https://doi.org/10.1007/s10578-011-0252-3.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. https://doi.org/10.1080/10705519909540118.

Hubbard, J. A., Parker, E. H., Ramsden, S. R., Flanagan, K. D., Relyea, N., Dearing, K. F., Smithmyer, C. M., Simons, R. F., & Hyde, C. T. (2004). The relations among observational, physiological, and self-report measures of children's anger. *Social Development, 13*, 14–39. https://doi.org/10.1111/j.1467-9507.2004.00255.x.

Katsurada, E., & Sugawara, A. I. (1998). The relationship between hostile attributional bias and aggressive behavior in preschoolers. *Early Childhood Research Quarterly, 13*, 623–636. https://doi.org/10.1016/S0885-2006(99)80064-7.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development, 71*, 107–118. https://doi.org/10.1111/1467-8624.00124.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*, 1198–1202. https://doi.org/10.1080/01621459.1988.10478722.

Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford.

Little, T. D., Brauner, J., Jones, S. M., Nock, M. K., & Hawley, P. H. (2003). Rethinking aggression: A typological examination of the functions of aggression. *Merrill-Palmer Quarterly, 49*, 343–369. https://doi.org/10.1353/mpq.2003.0014.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130–149. https://doi.org/10.1037/1082-989X.1.2.130.

Mathieson, L. C., Murray-Close, D., Crick, N. R., Woods, K. E., Zimmer-Gembeck, M., Geiger, T. C., & Morales, J. R. (2011). Hostile intent attributions and relational aggression: The moderating roles of emotional sensitivity, gender, and victimization. *Journal of Abnormal Child Psychology, 39*, 977–987. https://doi.org/10.1007/s10802-011-9515-5.

Murray-Close, D., & Ostrov, J. M. (2009). A longitudinal study of forms and functions of aggressive behavior in early childhood. *Child Development, 80*, 828–842. https://doi.org/10.1111/j.1467-8624.2009.01300.x.

Murray-Close, D., Ostrov, J. M., Nelson, D. A., Crick, N. R., & Coccaro, E. F. (2010). Proactive, reactive, and romantic relational aggression in adulthood: Measurement, predictive validity, gender differences, and associations with intermittent explosive disorder. *Journal of Psychiatric Research, 44*, 393–404. https://doi.org/10.1016/j.jpsychires.2009.09.005.

Murray-Close, D., Nelson, D. A., Ostrov, J. M., Casas, J. F., & Crick, N. R. (2016). Relational aggression: A developmental psychopathology perspective. In D. Cicchetti (Ed.), *Developmental psychopathology* (3rd ed., pp. 660–722). Hoboken: Wiley. https://doi.org/10.1002/9781119125556.devpsy413.

Muthén, L. K., & Muthén, B. O. (1998–2020). Mplus (version 8.4) [computer software]. Los Angeles: Author.

Nelson, D. A., Mitchell, C., & Yang, C. (2008). Intent attributions and aggression: A study of children and their parents. *Journal of Abnormal Child Psychology, 36*, 793–806. https://doi.org/10.1007/s10802-007-9211-7.

Orobio de Castro, B., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child Development, 73*, 916–934. https://doi.org/10.1111/1467-8624.00447.

Ostrov, J. M., & Crick, N. R. (2007). Forms and functions of aggression during early childhood: A short-term longitudinal study. *School Psychology Review, 36*(1), 22–43.

Ostrov, J. M., Kamper-DeMarco, K. E., Blakely-McClure, S. J., Perry, K. J., & Mutignani, L. (2019). Prospective associations between aggression/bullying and adjustment in preschool: Is general aggression different from bullying behavior? *Journal of Child and Family Studies, 28*(9), 2572–2585.

Ostrov, J. M., & Keating, C. F. (2004). Gender differences in preschool aggression during free play and structured interactions: An observational study. *Social Development, 13*, 255–277. https://doi.org/10.1111/j.1467-9507.2004.000266.x.

Ostrov, J. M., Murray-Close, D., Godleski, S. A., & Hart, E. J. (2013). Prospective associations between forms and functions of aggression and social and affective processes during early childhood. *Journal of Experimental Child Psychology, 116*, 19–36. https://doi.org/10.1016/J.jecp.2012.12.009.

Pellegrini, A. D. (2004). *Observing children in their natural worlds: A methodological primer* (2nd ed.). Mahwah: Erlbaum.

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *Journal of Personality Assessment, 87*, 102–112. https://doi.org/10.1207/s15327752jpa8701_09.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The children's behavior questionnaire. *Child Development, 72*, 1394–1408. https://doi.org/10.1111/1467-8624.00355.

Runions, K. C., & Keating, D. P. (2007). Young children's social information processing: Family antecedents and behavioral correlates. *Developmental Psychology, 43*, 838–849. https://doi.org/10.1037/0012-1649.43.4.838.

Runions, K. C., & Keating, D. P. (2010). Anger and inhibitory control as moderators of children's hostile attributions and aggression. *Journal of Applied Developmental Psychology, 31*, 370–378. https://doi.org/10.1016/j.appdev.2010.05.006.

Schultz, D., Groth, E., Vanderwalde, H., Shannon, K., Shuttlesworth, M., & Shanty, L. (2018). Assessment of hostile and benign intent attributions in early childhood: Can we elicit meaningful information? *Social Development, 27*, 401–414. https://doi.org/10.1111/sode.12274.

Song, J. H., Colasante, T., & Malti, T. (2020). Taming anger and trusting others: Roles of skin conductance, anger regulation, and trust in children's aggression. *British Journal of Developmental Psychology, 38*, 42–58. https://doi.org/10.1111/bjdp.12304.

Sroufe, L. A. (2013). The promise of developmental psychopathology: Past and present. *Development and Psychopathology, 25*, 1215–1224. https://doi.org/10.1017/S0954579413000576.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173–180. https://doi.org/10.1207/s15327906mbr2502_4.

Stoddard, J., Scelsa, V., & Hwang, S. (2019). Irritability and disruptive behavior disorders. In A. K. Roy, M. A. Brotman, & E. Leibenluft (Eds.), *Irritability in pediatric psychopathology* (pp. 109–214). New York: Oxford University Press.

Sukhodolsky, D. G., Smith, S. D., McCauley, S. A., Ibrahim, K., & Piasecka, J. B. (2016). Behavioral interventions for anger, irritability, and aggression in children and adolescents. *Journal of Child and Adolescent Psychopharmacology, 26*, 58–64. https://doi.org/10.1089/cap.2015.0120.

Tackett, J. L., Kushner, S. C., Herzhoff, K., Smack, A. J., & Reardon, K. W. (2014). Viewing relational aggression through multiple lenses: Temperament, personality, and personality pathology. *Development and Psychopathology, 26*, 863–877. https://doi.org/10.1017/S0954579414000443.

Underwood, M. K. (2003). The comity of modest manipulation, the importance of distinguishing among bad behaviors. *Merrill-Palmer Quarterly, 49*, 373–389. https://doi.org/10.1253/mpq.2003.0016.

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *The Journal of Experimental Education, 80*, 26–44. https://doi.org/10.1080/00220973.2010.531299.

Yoon, M., & Millsap, R. E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling: A Multidisciplinary Journal, 3*, 435–463. https://doi.org/10.1080/10705510701301677.

Ziv, Y. (2012). Exposure to violence, social information processing, and problem behavior in preschool children. *Aggressive Behavior, 38*, 429–441. https://doi.org/10.1002/ab.21452.