**Supplemental Material**

**Reporter Selection and Confirmatory Factor Analyses**

As outlined in our preregistration, observations or observer reports were the first-choice informant for assessments of aggression and parents were the first-choice informant for assessments of temperament. However, alternative informants were considered in the case of missing data, restricted range, and measures with low reliability or validity. Following prior work with this dataset (citation omitted), given the restricted range of behavioral observations, composite scores of forms and functions of aggression were computed averaging z-scores of school-based behavioral observations (i.e., “naturalistic observations”) and behavioral ratings (i.e., “Research Assistant (RA) reports”). For temperament constructs with multiple indices and reporters (emotion dysregulation; fearlessness; conscience), we tested measurement models using confirmatory factor analyses (CFA; Bollen, 1989) to assess the utility of latent constructs. Only observed variables with significant and substantial factor loadings (see Brown, 2006) on the latent factors were retained.

For behavioral emotion dysregulation, a preliminary CFA was run with five indicators of emotion dysregulation (teacher-reported low emotion regulation skills, emotional lability/negativity, Hubbard and CBQ anger, and negative emotionality); the two residual variances from the anger measures were allowed to correlate. Low emotion regulation was from the emotion regulation subscale of the Emotion Regulation Checklist (ERC; Shields & Cicchetti, 1997) and teachers responded on a 4-point Likert scale (*1 = Never, 4 = Almost always*) to eight items (e.g., “Displays appropriate negative emotions in response to hostile, aggressive, or intrusive acts by peers”). Teacher reports rather than parent reports were selected for two reasons. First, there was a high level of missingness of parent reports at Time 1 (*N* = 157 had data) relative to teacher reports (*N* = 293 had data). In addition, parent reports had marginal to unacceptable reliability. Our initial CFA indicated marginal model fit (CFI = .94, RMSEA = .19, SRMR = .05). Dropping the low emotion regulation skills indicator yielded a model with excellent model fit (CFI = 1.0, RMSEA = .000, SRMR = .003), with standardized factor loadings ranging from .71 to .92, *p*s < .001 (see Table S2). A multi-group CFA run with gender as the grouping variable indicated that factor loadings did not differ across gender (Wald Δχ2(4) = 5.51, *p* = .24). Thus, factor loadings were constrained across gender in multigroup analyses.

For the construct of behavioral fearlessness, we conducted a preliminary CFA in which T1 parent- and teacher-reports of temperamental fearfulness (reverse-coded) and daring served as indicators of a latent fearlessness variable. Despite substantial missing parent-reported data, parent- and teacher-reported indices of behavioral fearlessness were included to provide four potential indicators of the latent variable. However, difficulties with model convergence indicated that this model provided a poor fit to the data. Follow-up analyses were conducted with parent- and teacher-reported daring and fearlessness separately, with parent- and teacher-reported indicators constrained to be equal to achieve model identification. Both models provided excellent fit to the data (CFI = 1.0; RMSEA = 0.0; SRMR = 0.0 in both models), suggesting the fearlessness and daring should be treated separately in analyses. Further, as only 2 indicators were included, and inclusion of latent variables of these constructs in our more complex structural models resulted in difficulties with model estimation, we used manifest composites of fearlessness and daring, rather than latent variables to simplify final models.

Next, we ran a CFA in which T1 teacher-reported rule internalization, empathy, and prosocial behavior served as indicators of conscience. For prosocial behavior, the Preschool Social Behavior Scale-Teacher Form (PSBS-TF; Crick et al., 1997) was used. Teachers responded to four items (e.g., “This child is kind to peers”) on a 5-point scale from 1 (*never or almost never true*) to *5* (*always or almost always true*). The measure had good internal consistency (Cronbach’s 𝛼 = .88) in the present study. For conscience, teacher-reports were selected due to missing data from parent reports, the fact that parent report was only collected at one time point (and thus would prohibit hypothesized mediation tests), and the availability of 3 indicators for a potential conscience latent construct. This CFA provided excellent fit to the data (CFI = 1.0; RMSEA = 0.00; SRMR = 0.0). However, a single model including this latent conscience variable as well as manifest fearlessness and daring provided a poor fit to the data (CFI = .77; RMSEA = .23; SRMR = 0.9). Modification indices indicated that the three indicators of conscience had distinct associations with daring, such that rule internalization, but not the other two indicators, were associated with daring. Given the theoretical focus on rule internalization and empathy as a key mechanism linking fearlessness with aggression (e.g., Frick & Morris, 2004; Kagan, 1998; Kochanska, 1993), we used the manifest variables of rule internalization and empathy to test key study hypotheses. Further, although we had proposed to also include prosocial behavior as an indicator of latent conscience in our preregistration, based on these analyses, only Empathy and Rule Internalization manifest variables were included in final models; the inclusion of these variables is aligned with the approach taken by Cornell and Frick (2007) and parallels the parent version of the form (Kochanska et al., 1994).

**Physiological: Measurement and Covariates**

To calculate RSA, ECG data was converted to interbeat interval (IBI) data and cleaned using CardioEdit software (Brain-Body Center, University of Illinois-Chicago, 2007). In this procedure, the research assistants manually edited the IBI data to remove artifacts due to movement, room noise, and digitizing error. The research assistants were trained to reliability prior to proceeding with cleaning the data for this study. Following the data cleaning procedure, RSA estimates were calculated from the IBI data using the CardioBatch software (Brain-Body Center, University of Illinois-Chicago, 2007) and the time series method outlined by Porges (1985). The calculation settings consisted of 30-second epochs and frequency parameters of .24 to 1.04 Hz.

As outlined in the preregistration, respiration was evaluated as a control variable for RSA models. However, as respiration was not correlated with RSA in the present sample, *r* = .01, *p* = .97, it was not included as a covariate in final models. As outlined in our preregistration, BMI and room temperature were investigated as potential covariates given their potential association with SCL; however, as SCL was not significantly related to either, they were not included in final models.

**Missing Data, Attrition, and Power for Physiology Subsample Analyses**

Attrition was not associated with gender [χ2(1) = 0.01 - 0.02, *p*s = .89 - .94], age [*F*s(1, 294) = 0.24 – 0.73, *p*s = .34 - .63], race/ethnicity [χ2(4) = 0.50 - 0.88, *p*s = .93 - .97], or study cohort [χ2(3) = 3.96 – 4.17, *p*s = .24 - .27] at either time point. Lower SES tended to be associated with attrition at T2 [*t*(299) = 1.89, *p* = .06, *d* = .24] and significantly associated with attrition at T3 [*t*(299) = 2.12, *p* = .04, *d* = .26] using a proxy SES variable based on school code, consistent with prior work (citation omitted). Therefore, SES is included as a covariate in all structural models. Attrition at both T2 and T3 was also associated with lower teacher-reported daring at T1 [*t*s(290) = 2.35 – 2.74, *p*s = .006 - .02, *d*s = .30 - .35].

The majority of variables in the current models had minimal missing data within timepoint (1% - 5%), with two exceptions. Parent report on daring and fearlessness were missing for 48% of participants (*n*s = 156 - 157). Those with parent report data did not differ from those without on any demographic variables (i.e., gender, age, SES, or race/ethnicity) or any predictor or outcome variables in the model. Little’s (1988) Missing Completely at Random (MCAR) test demonstrated that the data were most likely missing completely at random [*χ*2(468) = 511.38, *p* = .08]. Additionally, only a subset of participants opted to complete the physiology lab visit (*n* = 93, 31.33%). As mentioned in the main text, this subsample did not differ from the full sample on demographics (i.e., gender, SES, or race/ethnicity) or any predictor or outcome variables with the exception that those with physiology data had significantly lower levels of proactive physical aggression at T1 [*t*(223.70) = -2.45, *p* = .02 *d* = -.28] and were older [t(213.17) = 2.25, *p* = .04, *d* = .26)], which was to be expected given the aforementioned study design that began the psychophysiology assessment only after the participant was 48 months old. Consistent with the study preregistration, missing data was accommodated using FIML; further, given the large amount of missing physiology data, these analyses were completed with just the physiology subsample. Given the reduction in expected sample size for the physiological subsample, we conducted power analyses to determine whether proposed analyses with this subsample were adequately powered. For main effects, analyses conducted in GPower 3.1.9.7 indicated that power to detect a small to medium effect (f2 = .09) with a sample size of 93 and 6 predictors was .82. For nested model comparisons, power was .86 to detect significant differences across nested models, assuming *df* of 14, a difference in RMSEA from .06 to .04, and a change in *df* of 1 (MacCallum et al., 2006). However, following Schoemann et al. (2017), with moderate (.30) associations between the independent variable and mediator, and mediator and the dependent variable, and the independent and dependent variable, respectively, an *N* of 104 was required to achieve power of .80. Given that our sample size was smaller than *N* = 104, we did not include mediation analyses in the manuscript.
**Stability of Forms and Functions of Aggression**

For our primary models, we first conducted a multigroup path analysis model with gender as the grouping variable in which T3 forms and functions of aggression were regressed onto T1 forms and functions of aggression. SES was included as a covariate. As shown in **Figure 1** and **Table S4**, proactive and reactive physical aggression were stable across the course of the study for boys and girls. Neither proactive nor reactive relational aggression were stable over time.

**Hostile Attribution and Rejection as Potential Mediators of Associations between Emotion Dysregulation and Aggression**

Both Frick and Morris (2004) and Lahey et al. (2008) suggest that angry and negative reactions to environmental stressors may be directly associated with aggressive responding, as a tendency to experience anger is likely to elicit aggressive responses to peer conflict. However, emotion dysregulation may also result in aggressive behavior through impairment in social information processing skills, such as hostile attribution biases (Crick & Dodge, 1994; Lemerise et al., 2005). Moreover, there is evidence that hostile attribution biases (HAB) may be differentially and uniquely related to reactive physical and relational aggression (see Martinelli et al., 2018). Studies that have explored both forms of aggression have shown that children and adolescents identified as relationally aggressive display HAB for ambiguous provocations of a relational nature (e.g., being left out of a peer group; Crick, 1995; Crick et al., 2002). In contrast, physically aggressive children have been found to display HAB for ambiguous provocations of an instrumental nature (e.g., damage to property; Crick et al., 2002). Moreover, there is evidence that HAB may be differentially and uniquely related to reactive physical and relational aggression (see Martinelli et al., 2018).

Temperament theories of aggression also identify impaired social relationships, including peer rejection, as an additional mechanism; specifically, rejection is argued to result in limited opportunities for socialization against aggression (Frick & Morris, 2004; Waldman et al., 2011). In fact, peer rejection is associated with both physical and relational aggression (e.g., Crick & Grotpeter, 1995; Ostrov et al., 2013; Zimmer-Gembeck et al., 2005). Interestingly, evidence indicates that rejection is more strongly related to reactive functions of aggression as youth may respond to peer dislike with aggressive conduct (e.g., Card & Little, 2006; Evans et al., 2019). Thus, we examined both proposed mediators (i.e., HAB and peer rejection). Based on the available theory and literature, we predicted that HAB for relational provocations would mediate the association between emotion dysregulation and reactive relational aggression whereas HAB for instrumental provocations would mediate the association between emotion dysregulation and reactive physical aggression. We also expected that peer rejection would mediate the associations between emotion dysregulation and reactive physical as well as reactive relational aggression.

**Peer Rejection (Time 2)**. Peer rejection was measured with the PSBS-TF (Crick et al., 1997; Ostrov et al., 2004), which assessed dislike by same and opposite sex peers. Teachers responded to two items (e.g., “This child is disliked by peers of the same sex”) on the aforementioned 5-point scale, which had good internal consistency (Cronbach’s 𝛼 = .92) in the present study. Final analyses were conducted using a weighted sum across the items.

**Child Interview – HAB-I and HAB-R.** Hostile attribution biases (HAB) were assessed at T2 through child interviews by trained graduate students using a modified version of the Assessment of Preschool Social Information Processing (APSIP; Casas & Crick, 1999), which was adapted (Godleski & Ostrov, 2020) to include cartoon faces depicting each of the response options. Eight provocation stories in which the intent of the provocateur was ambiguous were read to the child: four stories were instrumental (e.g., conflict over a resource) and four stories were relational (e.g., not receiving an invitation to a party) in nature. Children reported on whether the provocateur was “Not trying to be mean” (Score = 0), “Kind of trying to be mean” (Score = 1), or “Really trying to be mean” (Score = 2; see Godleski & Ostrov, 2020, for details). Responses to the relational and instrumental stories, respectively, are summed. Similar approaches have yielded acceptable reliability (e.g., Crick et al., 2002). Prior work has provided support for reliability of both HAB subtypes with one story removed (Cronbach’s αs = .69 - .72) and convergent validity between with parent reports (Godleski & Ostrov, 2020). In the current study, removing one vignette from each subscale improved reliability (Cronbach’s αs = .62 - .64 rather than αs = .56 - .57). Therefore, consistent with the preregistered plan, subscales with one vignette removed were retained for all models.

***Hostile Attribution and Rejection Mediation Analyses***

First, we tested if hostile attribution biases (HAB) for instrumental and relational provocations assessed at T2 mediated the relation between T1 emotion dysregulation and T3 reactive physical and reactive relational aggression, respectively, and if indirect effects were moderated by gender. As detailed in our preregistration, analyses were run separately by subtype of aggression; both types of HAB were included as simultaneous mediators in each model. Results indicated that indirect effects were not moderated by gender; further, there was no evidence of indirect effects for any form or function of aggression for boys or girls (see **Table S6**). Proposed follow-up robustness tests including all forms/functions of aggression resulted in model estimation difficulties and thus are not reported.

We also examined whether T2 rejection mediated the association between T1 emotion dysregulation and T3 reactive physical and reactive relational aggression, and whether these indirect effects were moderated by gender. All forms/functions of aggression were included in a single model. Multigroup path analyses run by gender indicated that indirect effects were not moderated by gender, and there were no significant indirect effects of emotion dysregulation at T1 to T3 forms/functions of aggression via rejection for boys or girls (see **Table S7**).

**Supplemental References**

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons. https://doi-org/10.1002/9781118619179

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press.

Cardio Batch Software, Brain-Body Center, University of Illinois at Chicago, 2007.

Cardio Edit Software, Brain-Body Center, University of Illinois at Chicago, 2007.

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

Cornell, A. H., & Frick, P. J. (2007). The moderating effects of parenting styles in the association between behavioral inhibition and parent reported guilt and empathy in preschool children. *Journal of Clinical Child and Adolescent Psychology, 36*(3), 305-318. https://doi.org/10.1080/15374410701444181

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., Casas, J. F., & Mosher, M. (1997). Relational and overt aggression in preschool. *Developmental Psychology, 33*(4), 579–588. https://doi-org/10.1037/0012-1649.33.4.579

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.

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

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*. https://doi-org/10.1016/j.jecp.2019.104775

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using

G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149-1160. http://dx.doi.org/10.3758/BRM.41.4.1149

Kagan, J. (1998). Biology and the child. In N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development., 3* (5), 177–235. John Wiley & Sons, Inc.

Kochanska, G., DeVet, K., Goldman, M., Murray, K., & Putnam, S. P. (1994). Maternal reports of conscience development and temperament in young children. *Child Development, 65*(3), 852-868.

Lemerise, E. A., Gregory, D. S., & Fredstrom, B. K. (2005). The influence of provocateurs’ emotion displays on the social information processing of children varying in social adjustment and age. *Journal of Experimental Child Psychology, 90,* 344-366.

MacCallum, R. C., Browne, M. W., & Cai, L. (2006). Testing differences between nested

covariance structure models: Power analysis and null hypotheses. *Psychological Methods, 11(1),* 19-35. http://dx.doi.org/10.1037/1082-989X.11.1.19

Martinelli, A., Ackermann, K., Bernhard, A., Freitag, C. M., & Schwenck, C. (2018). Hostile attribution bias and aggression in children and adolescents: A systematic literature review on the influence of aggression subtypes and gender. *Aggression and Violent Behavior, 39,* 25-32. https://doi.org/10.1016/j.avb.2018.01.005.

Ostrov, J. M., Woods, K. E., Jansen, E. A., Casas, J. F., & Crick, N. R. (2004). An observational study of delivered and received aggression, gender, and social-psychological adjustment in preschool: “This White Crayon Doesn’t Work*.” Early Childhood Research Quarterly, 19*(2), 355–371. https://doi-org/10.1016/j.ecresq.2004.04.009

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size

for simple and complex mediation models. *Social Psychological and Personality Science, 8(4)*, 379-386. http://dx.doi.org/10.1177/1948550617715068

Waldman, I. D., Tackett, J. L., Van Hulle, C. A., Applegate, B., Pardini, D., Frick, P. J., & Lahey, B. B. (2011). Child and adolescent conduct disorder substantially shares genetic influences with three socioemotional dispositions. *Journal of Abnormal Psychology, 120*(1), 57–70. https://doi-org/10.1037/a0021351

Zimmer-Gembeck, M. J., Geiger, T. C., & Crick, N. R. (2005). Relational and physical aggression, prosocial behavior, and peer relations: Gender moderation and bidirectional associations. *The Journal of Early Adolescence, 25*(4), 421–452. https://doi-org/10.1177/0272431605279841

**Table S1.**

*Gender Differences in Study Variables*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Boys | Girls |  |  |
|  | *Mean* | *SD* | *Mean* | *SD* | *F* | *p* |
| Pro Pagg T1 | .10 | .87 | -.09 | .74 | 2.38 | .09 |
| Rea Pagg T1 | .12 | .94 | -.12 | .81 | 3.02 | .05 |
| Pro Ragg T1 | -.04 | .80 | .04 | .78 | .45 | .64 |
| Rea Ragg T1 | -.08 | .70 | .11 | .93 | 1.97 | .14 |
| Pro Pagg T3 | .14 | .89 | -.18 | .60 | 4.71 | .01 |
| Rea Pagg T3 | .14 | .81 | -.20 | .62 | 5.93 | .003 |
| Pro Ragg T3 | -.05 | .78 | .09 | .75 | 1.01 | .37 |
| Rea Ragg T3 | -.08 | .71 | .07 | .83 | 1.41 | .25 |
| Emotion Dysreg LV T1 | .00 | 1.0 | -.18 | 1.0 | NA | .17 |
| Daring TR T1 | 2.39 | .91 | 2.05 | .71 | 5.73 | .004 |
| Fearless TR T1 | 4.66 | 1.40 | 4.50 | 1.30 | .54 | .59 |
| Int TR T1 | 4.07 | 1.69 | 4.69 | 1.55 | 5.24 | .006 |
| Empathy TR T1 | 4.31 | 1.40 | 4.63 | 1.06 | 2.82 | .06 |
| SCL T1 | 15.06 | 7.80 | 12.74 | 7.82 | .90 | .41 |
| RSA T1 | 6.85 | 1.33 | 7.04 | 1.21 | 1.82 | .17 |
| Daring PR T1 | 2.52 | .66 | 2.41 | .69 | .53 | .59 |
| Fearless PR T1 | 4.47 | 1.16 | 4.40 | 1.08 | .55 | .58 |
| Int TR T2 | 3.79 | 1.53 | 4.88 | 1.60 | 12.54 | .001 |
| Empathy TR T2 | 4.21 | 1.12 | 4.87 | 1.09 | 10.12 | .001 |

*Note*. Pro = Proactive, Rea = Reactive, Ragg = Relational aggression, Pagg = Physical aggression, TR = Teacher report, LV = Latent Variable, Dysreg = Dysregulation, Int = Internalization, SCL = Skin Conductance Level, RSA = Respiratory Sinus Arrythmia, T1 = Time 1, T2 = Time 2, T3 = Time 3. Descriptive statistics for the latent variable is not included. Descriptives for the latent variable were examined in Mplus, whereas descriptives for the observed variables were examined in SPSS.

**Table S2**

*Standardized Factor Loadings for Emotion Dysregulation Latent Factor*

|  |  |
| --- | --- |
| Item | Loading |
| Anger/Frustration TR | .71 |
| CADS Negative Emotionality TR | .84 |
| CBQ Anger TR | .78 |
| ERC Emotional Lability/Negativity TR | .92 |

*Note.* All loadings were significant at *p* < .001. TR = Teacher report, CADS = Child and Adolescent Dispositions Scale, CBQ = Child Behavior Questionnaire, ERC = Emotion Regulation Checklist. The residual variances from the Anger/Frustration and CBS anger scales were allowed to correlate.

|  |  |
| --- | --- |
| **Table S3***Item Level Correlations for Composite Indicators at Time 1* |  |
|  | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. | 10. | 11. | 12. |
| 1. CBQ TR- Anger |  |  |  |  |  |  |  |  |  |  |  |  |
| 2. CBQ TR-Fearlessness (R) |  -.29\*\* |  |  |  |  |  |  |  |  |  |  |  |
| 3. CADS TR- Neg emotionality | .65\*\* | -.29\*\* |  |  |  |  |  |  |  |  |  |  |
| 4. CADS TR- Daring |  .40\*\* |  -.12\* |  .32\*\* |  |  |  |  |  |  |  |  |  |
| 5. TR- Internalization | -.55\*\* |  -.04 | -.51\*\* | -.36\*\* |  |  |  |  |  |  |  |  |
| 6. TR- Empathy | -.11 | -.35\*\* | -.19\*\* |  .09 |  .38\*\* |  |  |  |  |  |  |  |
| 7. ERC TR- Emotion regulation | -.34\*\* | -.16\* | -.48\*\* |  .02 |  .26\*\* | .43\*\* |  |  |  |  |  |  |
| 8. ERC TR- Lability/ negativity |  .72\*\* | -.27\*\* |  .77\*\* |  .36\*\* | -.57\*\* | -.22\*\* | -.55\*\* |  |  |  |  |  |
| 9. TR- Anger/Frustration |  .67\*\* |  -.13\* |  .61\*\* |  .50\*\* | -.54\*\* | -.08 |  -.18\*\* | .66\*\* |  |  |  |  |
| 10. PSBS TR- Prosocial behavior | -.40\*\* |  -.12 | -.51\*\* | -.07 |  .60\*\* |  .58\*\* | .60\*\* | -.58\*\* | -.33\*\* | \_\_\_\_ |  |  |
| 11. CBQ PR-Fearlessness (R) |  -.04 |  .19\* |  -.15 | -.02 |  -.09 |  -.03 |  .05 |  -.02 |  -.04 | -.01 | \_\_\_\_\_ |  |
| 12. CADS PR- Daring |  .15 |  .10 |  .07 |  .29\*\* | -.23\*\* | -.03 | -.04 |  .07 |  .16\* | -.11 | .11 |  |
| *M* | 3.23 | 4.59 | 1.92 | 2.24 | 4.35 | 4.46 | 3.18 | 1.83 | 1.68 | 14.63 | 4.43 | 2.47 |
| *SD* | 1.60 | 1.36 | 0.74 | 0.84 | 1.66 | 1.27 | 0.51 | 0.53 | 0.58 | 3.04 | 1.13 | 0.67 |
| Range | 1.00-6.83 | 1.00-7.00 | 1.00-4.00 | 1.00-4.00 | 1.00-7.00 | 1.00-7.00 | 1.65-4.00 | 1.00-3.44 | 1.00-3.45 | 6.00-20.00 | 1.67-6.83 | 1.00-4.00 |

*Note.* \* *p* < .05, \*\**p* < .01. TR = Teacher report, PR = Parent report, Neg = Negative, CADS = Child and Adolescent Dispositions Scale, CBQ = Child Behavior Questionnaire, ERC = Emotion Regulation Checklist, PSBS = Preschool Social Behavior Scale.

**Table S4**

*Standardized Estimates for Associations in Stability Model*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **T1 Concurrent Associations** | T1 Proactive Physical | T1 Reactive Physical | T1 Proactive Relational | T1 Reactive Relational |
| T1 Proactive Physical |  |  |  |  |
| T1 Reactive Physical | .37\*\*\*/.43\*\*\* |  |  |  |
| T1 Proactive Relational | .10\*/.12\* | .12\*/.14\* |  |  |
| T1 Reactive Relational | *.08/.37\** | .15\*/.12\* | .24\*\*\*/.20\*\*\* |  |
| SES | .18\*/.19\* | .24\*\*/.26\*\* | .09/.10 | .16+/.12+ |
|  |  |  |  |  |
| **T3 Concurrent Associations** | T3 Proactive Physical | T3 Reactive Physical | T3 Proactive Relational |  |
| T3 Proactive Physical |  |  |  |  |
| T3 Reactive Physical | .61\*\*\*/.72\*\*\* |  |  |  |
| T3 Proactive Relational | .41\*\*\*/.49\*\*\* | .40\*\*\*/.43\*\*\* |  |  |
| T3 Reactive Relational | .34\*\*\*/.37\*\*\* | .31\*\*/.31\*\* | **.37\*\*/.55\*\*\*** |  |
|  |  |  |  |  |
|  |  |  |  |  |
| **Longitudinal Associations** | T3 Proactive Physical | T3 Reactive Physical | T3 Proactive Relational | T3 Reactive Relational |
| T1 Proactive Physical | .14\*/.15\* | .05/.05 | -.02/-.02 | .00/.00 |
| T1 Reactive Physical | *.31\*\*\*/.14+* | .21\*\*/.21\*\* | .14\*/.13\* | .09/.08 |
| T1 Proactive Relational | .04/.05 | *.13/-.05* | .03/.03 | .01/.00 |
| T1 Reactive Relational | .07/.10 | .12/.16 | .11/.14 | .07/.09 |
| SES | -.04/-.05 | .05/.06 | .06/.06 | .03/.03 |

+*p* < .10, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001.

*Note.* Estimates show boys on left, girls on right. Italicized effects were marginally different across boys and girls (*p* < .10) whereas bolded effects were significantly different across boys and girls (*p* < .05). T1 = Time 1, T3 = Time 3. CFI = 1.0, RMSEA = .00, SRMR = .07.

**Table S5**

*Standardized Estimates for Associations between Behavioral Measures and Aggression*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Proactive Physical | Reactive Physical | Proactive Relational | Reactive Relational |
| **Model 1: Behavioral Emotion Dysregulation** |  |  |  |  |
| *Concurrent Associations* |  |  |  |  |
| T1 Emotion Dysregulation with T1 Aggression | .19\*/.20\* | .22\*\*/.23\*\* | **.01/.28\*\*\*** | **.06/.27\*\*** |
|  |  |  |  |  |
| *Longitudinal Associations* |  |  |  |  |
| T1 Emotion Dysregulation 🡪 T3 Aggression | .21\*/.26\* | .24\*\*/.26\*\* | .19\*/.21\* | **-.04/.45\*\*\*** |
|  |  |  |  |  |
| **Model 2: Behavioral Fearlessness** |  |  |  |  |
| *Concurrent Associations* |  |  |  |  |
| T1 Daring with T1 Aggression | .13\*/.15\* | .21\*\*/.24\*\* | .12\*/.14\* | .17\*/.14\* |
| T1 Fearlessness with T1 Aggression | -.03/-.03 | -.02/-.03 | -.07/-.08 | -.15/-.12 |
| T1 Rule Internalization with T1 Aggression | -.19\*\*/-.20\*\* | -.22\*\*\*/-.23\*\*\* | *-.06/-.13\** | **.00/-.16***+* |
| T1 Empathy with T1 Aggression | -.03/-.03 | .06/.07 | **.12**+**/-.05** | *.23\*\*/.07* |
|  |  |  |  |  |
| *Longitudinal Associations* |  |  |  |  |
| T1 Daring 🡪 T3 Aggression | .22\*\*\*/.25\*\*\* | *.32\*\*\*/.09* | .06/.06 | .06/.05 |
| T1 Fearlessness 🡪 T3 Aggression | -.07/-.08 | -.15*+*/-.16*+* | -.03/-.03 | -.09/-.07 |
| T1 Rule Internalization 🡪 T3 Aggression | -.07/-.08 | -.04/-.05 | **-.17/-.33\*\*** | **.07/-.40\*\*\*** |
| T1 Empathy 🡪 T3 Aggression | .03/.04 | -.01/-.01 | .20\*\*/.18\*\* | .06/.05 |

+*p* < .10, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001.

*Note*. Estimates show boys on left, girls on right. Italicized effects were marginally different across boys and girls (*p* < .10) whereas bolded effects were significantly different across boys and girls (*p* < .05). T1 = Time 1, T3 = Time 3. Model 1: CFI = .97, RMSEA = .05, SRMR = .07; Model 2: CFI = 1.0, RMSEA = .00, SRMR = .07.

**Table S6**

*Test of Indirect Effects of Behavioral Emotion Dysregulation to Aggression via Hostile Attributions*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Outcome Variable** | **Indirect Effect** | **Estimate** | **SE** | ***t*** | ***p*** |
| 1 | Proactive Physical | HAB-I | .009/.003 | .017/.012 | .53/.25 | .61/.80 |
|  |  | HAB-R | .004/.007 | .026/.021 | .17/.36 | .87/.74 |
|  |  | *Total Indirect* | .013/.010 | .027/.024 | .50/.43 | .62/.69 |
|  |  |  |  |  |  |  |
| 2 | Reactive Physical | HAB-I | .017/-.002 | .021/.012 | .81/-.15 | .43/.89 |
|  |  | HAB-R | .000/.001 | .014/.015 | -.01/.05 | .99/.96 |
|  |  | *Total Indirect* | .017/-.001 | .022/.020 | .78/-.05 | .44/.96 |
|  |  |  |  |  |  |  |
| 3 | Proactive Relational | HAB-I | .005/.011 | .018/.020 | .25/.53 | .80/.62 |
|  |  | HAB-R | .005/-.001 | .024/.014 | .20/-.05 | .85/.97 |
|  |  | *Total Indirect* | .009/.010 | .027/.025 | .35/.40 | .73/.71 |
|  |  |  |  |  |  |  |
| 4 | Reactive Relational | HAB-I | .008/.010 | .018/.024 | .44/.43 | .66/.67 |
|  |  | HAB-R | .005/-.002 | .018/.015 | .27/-.13 | .79/.89 |
|  |  | *Total Indirect* | .013/.008 | .021/.029 | .61/.29 | .54/.77 |

*Note*. Estimates show boys on left, girls on right. Gender did not moderate any indirect effects.

**Table S7**

*Test of Indirect Effects of Behavioral Emotion Dysregulation to Aggression via Rejection*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Outcome Variable** | **Estimate** | **SE** | ***t*** | ***p*** |
| Proactive Physical | -.028/.028 | .039/.051 | -.72/.56 | .47/.61 |
| Reactive Physical | -.010/.038 | .040/.057 | -.24/.67 | .81/.51 |
| Proactive Relational | .018/.038 | .052/.040 | .34/.97 | .74/.37 |
| Reactive Relational | .031/.037 | .041/.045 | .75/.82 | .45/.45 |

*Note*. Estimates show boys on left, girls on right. Gender did not moderate any indirect effects.

**Table S8**

*Test of Indirect Effects of Behavioral Fearlessness and Daring to Aggression via Rule Internalization and Empathy*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Outcome Variable** | **Predictor** | **Indirect Effect** | **Estimate** | **SE** | ***t*** | ***p*** |
| Proactive Physical | Daring | Rule Internalization | .018/.063 | .024/.051 | .75/1.23 | .47/.25 |
|  |  | Empathy | -.020/.006 | .018/.021 | -1.11/.28 | .28/.79 |
|  |  | *Total Indirect* | -.002/.069 | .026/.058 | -.08/1.19 | .94/.27 |
|  | Fearlessness | Rule Internalization | .011/.022 | .021/.029 | .53/.76 | .61/.49 |
|  |  | Empathy | .003/.007 | .022/.025 | .15/.27 | .89/.80 |
|  |  | *Total Indirect* | .014/.029 | .025/.041 | .56/.70 | .58/.53 |
|  |  |  |  |  |  |  |
| Reactive Physical | Daring | Rule Internalization | .016/.030 | .022/.052 | .71/.58 | .48/.58 |
|  |  | Empathy | -.002/.007 | .018/.021 | -.13/.33 | .90/.74 |
|  |  | *Total Indirect* | .014/.037 | .022/.057 | .62/.65 | .54/.53 |
|  | Fearlessness | Rule Internalization | .010/.011 | .020/.022 | .48/.48 | .64/.66 |
|  |  | Empathy | .000/.008 | .013/.026 | .03/.31 | .98/.76 |
|  |  | *Total Indirect* | .010/.019 | .021/.036 | .47/.53 | .64/.62 |
|  |  |  |  |  |  |  |
| Proactive Relational | Daring | Rule Internalization | **.006/.132** | **.023/.057** | **.27/2.30** | **.79/.03** |
|  |  | Empathy | .019/.000 | .021/.012 | .90/-.04 | .39/.97 |
|  |  | *Total Indirect* | **.025/.131** | **.027/.059** | **.94/2.23** | **.36/.03** |
|  | Fearlessness | Rule Internalization | .004/.046 | .019/.053 | .19/.88 | .84/.40 |
|  |  | Empathy | -.003/-.001 | .019/.014 | -.16/-.04 | .88/.97 |
|  |  | *Total Indirect* | .001/.046 | .025/.054 | .03/.85 | .98/.41 |
| Reactive Relational | Daring | Rule Internalization | *-.010/.101* | *.032/.060* | *-.31/1.70* | *.75/.11* |
|  |  | Empathy | .004/-.004 | .016/.019 | .27/-.23 | .78/.82 |
|  |  | *Total Indirect* | -.006/.097 | .031/.060 | -.19/1.60 | .85/.13 |
|  | Fearlessness | Rule Internalization | -.006/.036 | .026/.046 | -.23/.77 | .81/.48 |
|  |  | Empathy | -.001/-.005 | .011/.025 | -.06/-.21 | .95/.84 |
|  |  | *Total Indirect* | -.007/.030 | .026/.045 | -.26/.67 | .79/.53 |

*Note*. Estimates show boys on left, girls on right. Italicized indirect effects were marginally different across boys and girls (*p* < .10) whereas bolded indirect effects were significantly different across boys and girls (*p* < .05).

**Table S9**

*Standardized Estimates for Associations in Stability Model for Physiology Subsample*

|  |  |
| --- | --- |
|  | Form of Aggression |
|  | Physical Aggression Model | Relational Aggression Model |
| **T1 Concurrent Associations** | T1 Proactive  | T1 Reactive | T1 Proactive | T1 Reactive |
| T1 Reactive Aggression | .33+/.47+ |  | .52\*\*/.28\*\* |  |
| SES | .19\*\*/.34\*\* | .37\*\*/.37\*\* | .10/.08 | .18+/.16+ |
| BMI | **-.23\*/.15** | *-.35\*\*\*/.01* | **-.03/.30\*** | **-.21\*\*/.23\*** |
|  |  |  |  |  |
| **T3 Concurrent Associations** | T3 Proactive |  | T3 Proactive |  |
| T3 Reactive Aggression | .41\*\*/.71\*\* |  | .47\*\*\*/.47\*\*\* |  |
|  |  |  |  |  |
| **Longitudinal Associations** | T3 Proactive | T3 Reactive | T3 Proactive | T3 Reactive |
| T1 Proactive Aggression | .22\*/.18\* | -.15/-.11 | **.51\*\*/.03** | .10/.12 |
| T1 Reactive Aggression | .18+/.27+ | .35\*\*/.46\*\* | **-.47\*\*\*/-.05** | -.12/-.12 |
| SES | -.04/-.07 | .08/.11 | .20\*/.32\* | -.14/-.12 |
| BMI | .06/.07 | -.11/-.13 | -.14/-.19 | -.16/-.11 |

+*p* < .10, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001.

*Note*. Models were run separately by form of aggression due to estimation difficulties including all aggression subtypes in a single model with the physiology subsample. Estimates show boys on left, girls on right. Italicized effects were marginally different across boys and girls (*p* < .10) whereas bolded effects were significantly different across boys and girls (*p* < .05). T1 = Time 1, T3 = Time 3. Physical aggression model: CFI = .98, RMSEA = .06, SRMR = .13; Relational aggression model: CFI = 1.0, RMSEA = .00, SRMR = .06.

**Table S10**

*Standardized Estimates for Associations between Physiology Measures and Aggression*

|  |  |
| --- | --- |
|  | Form of Aggression |
|  | Physical Aggression Model | Relational Aggression Model |
|  | Proactive | Reactive | Proactive | Reactive |
| **RSA** |  |  |  |  |
| *Concurrent Associations* |  |  |  |  |
| T1 RSA with T1 Aggression | *.12/-.31+* | **.04/-.55\*\*** | -.19+/-.13+ | .02/.01 |
|  |  |  |  |  |
| *Longitudinal Associations* |  |  |  |  |
| T1 RSA 🡪 T3 Aggression | **.28\*/-.07** | .07/.09 | -.08/-.11 | .22+/.14+ |
|  |  |  |  |  |
|  |  |  |  |  |
| **SCL** |  |  |  |  |
| *Concurrent Associations* |  |  |  |  |
| T1 SCL with T1 Aggression | -.02/-.03 | -.07/-.07 | -.03/-.02 | .03/.02 |
| T1 Rule Internalization with T1 Aggression | -.05/-.09 | -.25\*/-.25\* | -.21+/-.13+ | -.07/-.05 |
| T1 Empathy with T1 Aggression | **.25/-.24** | .07/.09 | *-.02/-.18* | .05/.04 |
|  |  |  |  |  |
| *Longitudinal Associations* |  |  |  |  |
| T1 SCL 🡪 T3 Aggression | -.06/-.09 | -.16\*/-.20\* | -.11/-.17 | -.04/-.03 |
| T1 Rule Internalization 🡪 T3 Aggression | -.21+/-.30+ | -.05/-.06 | -.40\*\*\*/-.59\*\*\* | -.46\*\*/-.36\*\* |
| T1 Empathy 🡪 T3 Aggression | -.01/-.01 | -.10/-.10 | **.03/.49\*\*\*** | .30/.18 |

+*p* < .10, \**p* < .05, \*\**p* < .01, \*\*\**p* < .001.

*Note*. Models were run separately by form of aggression. Estimates show boys on left, girls on right. Italicized effects were marginally different across boys and girls (*p* < .10) whereas bolded effects were significantly different across boys and girls (*p* < .05). RSA = Respiratory Sinus Arrhythmia; SCL – Skin Conductance Level, T1 = Time 1, T3 = Time 3. RSA Physical Aggression Model: CFI = .98, RMSEA = .05, SRMR = .12; RSA Relational Aggression Model: CFI = 1.0, RMSEA = .00, SRMR = .06; SCL Physical Aggression Model: CFI = 1.0, RMSEA = .00, SRMR = .10; SCL Relational Aggression Model: CFI = 1.0, RMSEA = .00, SRMR = .07.

**Figure S1**

*Stability Models in the Physiological Subsample*

1. ***Physical Aggression Stability Model***

****

1. ***Relational Aggression Stability Model***



*Note*. †*p* < .10, \**p* < .05, \*\**p* < .01. Agg = Aggression, SES = Socioeconomic status, BMI = Body Mass Index, T1 = Time 1, T3 = Time 3. Path estimates show boys on left, girls on right. Bolded estimates indicate a significant difference between boys and girls. All non-bolded paths were constrained to be equal across gender but estimates may differ slightly due to differences in standard errors. Only significant paths are shown but all autoregressive and covariance paths were estimated. SES and BMI were included as covariates.