Peer Rejection and Academic Performance in the Link between Aggression and Depressive Symptoms: A Longitudinal Examination of Alternative Developmental Pathways

By

Spencer C. Evans, M.A

Submitted to the graduate degree program in Clinical Child Psychology in the Departments of Psychology and Applied Behavioral Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Paula J. Fite, Ph.D., Chair

_____

Michael C. Roberts, Ph.D., ABPP

_____

Eric M. Vernberg, Ph.D., ABPP

_____

Christopher C. Cushing, Ph.D.

_____

Anne P. Williford, Ph.D.

Date Defended: June 14, 2016

The Dissertation Committee for Spencer C. Evans
certifies that this is the approved version of the following dissertation:

Peer Rejection and Academic Performance in the Link between Aggression and Depressive Symptoms: A Longitudinal Examination of Alternative Developmental Pathways

_____
Paula J. Fite, Ph.D., Chair

Date Approved: June 14, 2016

Abstract

Background: The failure model posits that peer rejection and poor academic performance may account for the longitudinal association between children's aggressive behavior and subsequent depressive symptoms. However, the theory is unclear and evidence is mixed regarding the temporal sequence, relative magnitude, and possible interaction of these two mediators. Incorporating the functions of aggressive behavior may shed further light on these developmental pathways given that reactive, but not proactive, aggression is particularly associated with depressive symptoms and poor social and academic functioning. The present study investigated alternative pathways from early reactive aggression to subsequent peer rejection, academic performance, and depressive symptoms.

Method: A school sample of children (N = 963; ages 6-12; 49% female) was assessed annually by their primary classroom teachers over three years. Ratings of proactive and reactive aggression, peer rejection, academic performance, and depressive symptoms were collected. Following an accelerated design, path models were estimated to examine peer rejection and academic performance as developmental pathways from aggression subtypes in first grade to depressive symptoms in fifth grade. Developmental cascade sequences, interactions, and gender differences were also examined.

Results: Reactive aggression in first grade predicted peer rejection and poor academic performance in third grade, whereas proactive aggression predicted better academic performance. For girls, proactive aggression predicted lower peer rejection, and the path from reactive aggression to peer rejection was stronger than it was for boys. From third to fifth grade, peer rejection predicted subsequent depressive symptoms for boys only. The direct and moderated academic pathways to depressive symptoms were nonsignificant.

Conclusions: Results provide partial support for the failure model, particularly for the developmental pathways from reactive aggression to peer rejection to depressive symptoms. These findings highlight the importance of reactive aggression and peer functioning as key developmental mechanisms during middle childhood and as possible targets for assessment and intervention in the school context. Early reactive aggression appears to be a central risk factor for social and academic problems, while peer rejection in third grade appears to be a risk factor for depressive symptoms for boys. Limitations and implications for future research are discussed.

Keywords: failure model, proactive and reactive aggression, peer rejection, academic performance, depressive symptoms, developmental pathways

Acknowledgments

This dissertation would not have been possible without the help of many individuals. In particular, I am grateful to Paula for her consistent support, guidance, and patience in the development of this and many other research projects. I'd also like to thank Michael for his continued mentorship over the years. I am extremely fortunate to have not one but two outstanding mentors who have made invaluable contributions to my professional and personal development throughout graduate school. I also greatly appreciate the thoughtful comments and questions from the rest of my committee members, Eric, Christopher, and Anne. This dissertation is a stronger document as a result of your feedback.

In terms of the resources and partnerships that made this study possible, I am grateful to the elementary school administrators and staff for their continued partnership over the years. I would also like to thank all of the teachers who have participated and provided such valuable data. Of course, the research partnership is a result of Paula's efforts and I deeply appreciate that groundwork and the initial waves of data that made the rest of this dissertation possible. With respect to funding, I gratefully acknowledge the support of the University of Kansas (Lillan Jacobey Baur Early Childhood Fellowship; Doctoral Student Research Fund ×2; CCPP Pioneers Classes Dissertation Research Award), as well as the American Psychological Foundation (Elizabeth Munsterberg Koppitz Child Psychology Graduate Fellowship).

Surviving graduate school and completing a dissertation would not have been possible without the support of peers, friends, and family members. You have made the last five years not only bearable but even somewhat enjoyable! Thanks to all of the great people in Michael's lab, Paula's lab, and the clinical child program as a whole—too many people to name. I am lucky to be a part of this community. Patrick, thanks for being an awesome friend, and for the countless

hours of free stats consultation. Ashley, I can't tell you how much your support and encouragement has meant to me over this past year and a half. We make a great team. Finally, I am particularly grateful to my family. Mom and Dad, thanks for your continued support and encouragement throughout my entire education. Garrett and Calyssa, you guys are okay, too.

There are many, many more people who have had a tremendous impact on my life and helped me to get here. You know who you are. Thank you.

Table of Contents

List of Figures

List of Tables

Peer Rejection and Academic Performance in the Link between Aggression and Depressive Symptoms: A Longitudinal Examination of Alternative Developmental Pathways

Children who exhibit aggressive and antisocial behavior are at a greater risk for depressed mood in later childhood and adolescence (e.g., Capaldi, 1992; Panak & Garber, 1992; Van der Giessen et al., 2013), but less is known about the mechanisms that could account for this heterotypic transition. Consistent with a developmental psychopathology framework (Cicchetti & Rogosch, 2002), several possible models have emerged to support the multifinality of pathways emerging from aggression during childhood and the equifinality of pathways leading to depressive symptomatology in late childhood and adolescence. Broadly, these models include identifying the differences in the types of aggressive behavior exhibited in childhood (i.e., subgroups or subtypes) and exploring possible mechanisms that could help explain how these behaviors contribute to subsequent depressive symptoms (i.e., mediators or moderators).

Of note, Patterson and colleagues' "failure model" posits that aggressive children are more likely to encounter failures in two major domains of functioning—peer relationships and academic performance—and these failures, in turn, contribute to depressed mood over time (Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991). Despite its theoretical influence, this model has received little empirical evaluation from a longitudinal perspective. Findings have been mixed with regard to the relative contributions and temporal ordering of social and academic functioning as mediators. In addition, further insight can be gained by dividing aggression according to its functions (proactive and reactive), which demonstrate differential correlates and developmental sequelae in areas related to social, academic, and internalizing symptoms (Fite, Rathert, Colder, Lochman, & Wells, 2012a; Vitaro & Brendgen, 2012).

This study examines the roles of academic performance and peer rejection in the developmental pathways from reactive aggression to depressive symptoms among school-age children. Results advance the literature by (a) providing a longitudinal evaluation of four alternative mediation models, (b) distinguishing the developmental pathways based on the proactive and reactive functions of aggressive behavior, and (c) clarifying whether these pathways differ between boys and girls.

Developmental Pathways from Aggression to Depressed Mood

Children who exhibit significant aggressive behavior are at an increased risk for poorer outcomes in several areas of psychosocial functioning, including academic, social, and behavioral outcomes (Brennan, Shaw, Dishion, & Wilson, 2012; Campbell, Spieker, Burchinal, & Poe, 2006; Chen, Huang, Wang, & Li, 2010; Crick, Ostrov, & Werner, 2006; Vitaro & Brengden, 2012). Indeed, there is ample evidence that the strongest predictors of antisocial behavior in adolescence and adulthood include early and persistent patterns of such behaviors that emerge during childhood (Frick & Viding, 2009; Huesmann, Dubow, & Boxer, 2009; Moffitt, 2007). At the same time, children who exhibit significant aggressive and externalizing behaviors are also more likely to experience depression and internalizing problems, both concurrently and prospectively in childhood and adolescence (Campbell et al., 2006; Coie, Terry, Lenox, & Lochman, 1995; Panak & Garber, 1992) and even into adulthood (Loth, Drabick, Leibenluft, & Hulvershorn, 2014).

To help explain this developmental phenomenon, Patterson and colleagues (Granic & Patterson, 2006; Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991; Patterson, DeBaryshe, & Ramsey, 1989; Patterson & Yoerger, 1993) have developed, and continually refined, a unifying framework for understanding the causes, correlates, and outcomes of

antisocial behavior. According to this model, there are several dynamic, reciprocal processes that unfold in a cascading manner across development (Granic & Patterson, 2006; Patterson, DeBaryshe, & Ramsey, 1989). The model postulates that during early childhood the interplay among genetic, neurocognitive, and temperamental vulnerabilities, contextual risk factors, and parenting practices contribute to children's aggressive and antisocial behavior. In middle childhood, children who are aggressive and antisocial, while still possessing the associated risk factors noted above, are more likely to encounter psychosocial "failure" in multiple domains of functioning, including poor academic performance at school and rejection from the primary peer group. In turn, the cumulative wear and tear of these failures can contribute to depression and delinquency in later childhood and adolescence (Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991; Patterson & Yoerger, 1993).

The failure model has been highly influential in the developmental psychopathology and clinical child psychology literature. Subsequent extensions and variations of this model have been developed to explain, and at least partially support, an array of developmental phenomena, including the compounding effects of conduct problems *and* depressed mood (Capaldi, 1992; Ingoldsby et al., 2006), the development of depression in adulthood (Firth, Shohet, & Thurber, 2000), cannabis use disorders among adolescents (Marmorstein & Iacono, 2011), gender differences in psychosocial contributors to depression among adolescents (Boots, Wareham, & Weir, 2011), and the role of peer rejection and irritability in the development of internalizing problems and callous-unemotional traits (Barker & Salekin, 2012). Meanwhile, much of the subsequent work by Patterson and colleagues has built upon this model to focus on outcomes related to delinquency, crime, and antisocial behavior later in development, with relatively less

attention to the intermediary associations between peer rejection, academic performance, and depressed mood (e.g., Dishion, Véronneau, & Myers, 2010; Granic & Patterson, 2006).

Thus, it seems that much of the evidence related to the failure model has accumulated with relatively little rigorous testing of the original model. Indeed, the early articulations of the failure model (Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991; Patterson, DeBaryshe, & Ramsey, 1989) were largely based on a synthesis of empirical research and developmental theory from various sources. The model is useful in that it is amenable to empirical validation; however, it is difficult to test because it includes multiple processes that unfold simultaneously and sequentially across several domains of functioning (family, academics, peers) and over the course of child development.

Patterson and colleagues' early work was among the few attempts at testing their model directly. In a cross-sectional study with two cohorts of fourth-grade boys, Patterson and Capaldi (1990) examined the degree to which lower scores in peer relations, academic skills, and self-esteem were associated with concurrent depressive symptoms. The associations among peer rejection, academic skills, and depressed mood were all positive and significant in the measurement model (*r*s = .49 – .62). However, better model fit was obtained for a post hoc mediational model, where peer rejection was retained as the only direct predictor of depression, and peer rejection was modeled as a mediator, statistically accounting for the association between academic performance and depression.

Replications of this mediational model were subsequently tested among three cross-sectional samples of boys, including two cohorts from the previous study and a third at-risk sample (Patterson & Stoolmiller, 1991). Results supported the association between poor peer relations and depressed mood in all three samples (*r*s = .66 – .79), but the association between

poor academic skills and depressed mood was supported in only two of the three samples (*r*s = .53 – .67). The authors interpreted this as evidence that the social pathway was stronger and more generalizable than was the academic pathway; however, an alternative explanation also seems plausible. The sample for which the association between academic performance and depressed mood was nonsignificant consisted of 9- to 12-year-old boys whose parents had separated within 3-12 months prior to their participation in the study. One of the major risk factors of parental separation is the loss of close relationships between the child and the noncustodial parent, and this process is thought to be more common and more problematic for boys experiencing disrupted relationships with their fathers (Kelly & Emery, 2003). Thus, it seems likely that, compared to the other two samples, the boys in the separation sample were already primed for depressed mood by way of interpersonal rejection—or, at least, to a greater extent than by academic performance. In other words, Patterson and Stoolmiller's (1991) concerns about generalizability appear to be artifacts of the sample rather than the model, and the results should not be interpreted as a disconfirmation of the academic pathway. Moreover, the interpretations and generalizability of the findings from both studies are limited by their all-male samples and by their cross-sectional design.

Despite the persistent influence of the failure model, it seems that the notion of *dual* pathways has largely been cast aside in favor of models focusing on either social or (less commonly) academic problems as risk factors for depressed mood—or, alternatively, how these social and academic constructs relate to one another. Yet, evidence supporting the component associations within the failure model continues to emerge. For example, in a five-year longitudinal study among Chinese children, aggression had significant direct and indirect effects on subsequent social competence and academic achievement; however, the converse was not true

(Chen et al., 2010). Such developmental pathways appear to emerge relatively early, with similar evidence for aggression in toddlerhood predicting academic and social functioning in middle childhood (Brennan et al., 2012; Campbell et al., 2006) and showing incremental validity in accounting for later academic performance above and beyond personality factors (Barthelemy & Lounsbury, 2009). Other studies provide evidence for the second path in the mediation process, indicating that peer rejection (e.g., Gooren et al., 2011; Morrow et al., 2006), academic problems (e.g., McCarty et al., 2008), or both (e.g., Nocentini, Calamai, & Menesini, 2012; Obradović, Burt, & Masten, 2010) are uniquely associated with depressive symptoms. However, research on the nature and direction of the relationship between aggression and depressive symptoms is mixed. While meta-analyses and reviews show consistent evidence for longitudinal transitions from externalizing to internalizing problems (Burke, Loeber, Lahey, & Rathouz, 2005; Loth et al., 2014), the evidence is less clear regarding the role of aggression within this framework, in part because of the various ways in which aggression has been measured (e.g., physical aggression, antisocial behavior, externalizing problems). Here, evidence supports the utility of the functional subtypes of proactive and reactive aggression for elucidating these differential relations with psychosocial outcomes (Fite et al., 2012a; Vitaro & Brendgen, 2012).

Proactive and Reactive Functions of Aggression

Although aggression is often thought of as a monothetic construct, meaningful distinctions can be drawn according to the forms (e.g., relational vs. physical) and functions (reactive vs. proactive) of aggressive behavior (Little, Henrich, Jones, & Hawley, 2003; Vitaro & Brendgen, 2012). Of relevance to the present study are the function-based subtypes of proactive and reactive aggression (Dodge & Coie, 1987). Proactive aggression consists of goal-oriented, instrumental behaviors committed as a means of accomplishing a goal. By contrast, reactive

aggression consists of emotionally driven, impulsive behaviors that occur in response to a real or perceived threat (Fite et al., 2012a; Vitaro & Brendgen, 2012). Proactive aggression is best explained from a social learning perspective (Bandura, 1973), where children develop aggressive behavior through observation and imitation of aggressive modeling in their environment, and natural reinforcement of the aggressive behaviors. By contrast, reactive aggression is consistent with the frustration-aggression model (Berkowitz, 1989), in which an individual exhibits aggressive behavior in response to frustration or threat.

These two subtypes of aggression are theoretically distinct but empirically correlated, with estimates (*r*) typically between .6 and .8 in studies utilizing questionnaires (Card & Little, 2006; Polman, de Castro, Koops, van Boxtel, & Merk, 2007). Yet, factor analytic studies support the proactive-reactive dichotomy (Fite, Colder, & Pelham, 2010; Little et al., 2003; Raine et al., 2006). After controlling for their shared variance, proactive and reactive aggression demonstrate differential correlates, etiologies, and outcomes across development, including distinct associations with academic, social, and emotional constructs (Card & Little, 2006; Fite et al., 2012a; Vitaro & Brendgen, 2012; Vitaro, Brendgen, & Tremblay, 2002).

Reactive aggression and depressive symptoms. Several cross-sectional studies support the link between reactive aggression and depressed mood in children. Boys who are primarily reactively aggressive are rated as being less happy than their peers (Day, Bream, & Pal, 1992). Similarly, Mathieson and Crick (2010) found that, of four form × function subtypes of aggression, only reactive-relational aggression was uniquely associated with internalizing symptoms cross-sectionally, whereas proactive-relational aggression actually predicted a decrease in internalizing symptoms over a one-year interval. In a child psychiatric inpatient sample, reactive, but not proactive, aggression was uniquely and positively associated with

depressive symptoms and suicidal behavior (Fite, Stoppelbein, & Greening, 2009). With regard to longitudinal studies, Vitaro, Brendgen, and Tremblay (2002) found that, in a large, community sample, children who were classified as reactively aggressive showed significantly greater depressive symptoms at age 13 compared to those who were proactively aggressive or nonaggressive. Similarly, Fite et al. (2014) found that high reactive, but not proactive, aggression in adolescent males was a risk factor for the development of depressive symptoms three years later. Studies examining mediators and moderators of the link between reactive aggression and depressive symptoms (discussed below) provide further support for this association while also elucidating possible developmental mechanisms.

Reactive aggression and academic and social functioning. There is also strong evidence that reactive, but not proactive, aggression is associated with peer rejection and victimization (e.g., Evans et al., 2015; Renouf et al., 2010; Salmivalli & Helteenvuori, 2007; White & Kistner, 2011). A meta-analysis of 36 correlational studies found that, after controlling for proactive aggression, reactive aggression is differentially associated with less favorable ratings in social preference, peer acceptance, peer rejection, peer victimization, and internalizing problems (Card & Little, 2006). Similarly, there is limited but consistent evidence indicating that reactive, not proactive, aggression is uniquely associated with poor academic performance (Fite et al., 2013; Day et al., 1992). In explaining the negative link between the more general construct of aggression and academic performance (e.g., Chen et al., 2010; Wentzel & Asher, 1992), evidence suggests that it is the presence of reactive aggression that specifically contributes most to children's academic competence in elementary school (Day et al., 1992). More specific tests of failure and cascade models (discussed below) provide further support for the link between reactive aggression and peer rejection and poor academic performance.

Academic and Social Functioning and Depressive Symptoms

Cross-sectional evidence indicates that depressive symptoms are associated with both peer rejection (e.g., Fite et al., 2012b; Morrow et al., 2006, 2008) and poor academic performance (e.g., Fite et al., 2013; Nocentini et al., 2012). Further, peer rejection statistically accounts for the association between aggression and depressive symptoms (Fite et al., 2013; Morrow et al., 2006, 2008), and this association has been supported and further elucidated by longitudinal research. For example, Panak and Garber (1992) found that from third through fifth grade, aggression predicted subsequent depression, with peer rejection accounting for this association. Similarly, peer rejection consistently emerges as a mediator in the link between externalizing/conduct problems in kindergarten-age children and internalizing/depressive symptoms later in middle childhood (Gooren et al., 2011; van Lier & Koot, 2010) and early adolescence (Pederson et al., 2007). With respect to academic functioning, low GPAs predicted subsequent depressive symptoms among samples comprised of school-age children (boys and girls) over the course of two years (Schwartz, Goman, Duong, & Nakamoto, 2008). Moreover, academic failure predicted subsequent depressive episodes for adolescent girls but not boys (McCarty et al., 2008).

Variations on a Theme: Evidence for Alternative Mediation Models

The literature reviewed above provides support for the specific components of the failure model (i.e., pathways from aggression to depressive symptoms, from aggression to social/academic functioning, and from social/academic functioning to depressive symptoms), while also implicating *reactive* aggression as the key facet of externalizing and antisocial behavior that conveys risk for those developmental trajectories. However, there are only a few longitudinal studies that have examined anything resembling the entire failure model (i.e., all

four variables over time), and none of these have specifically considered the roles of proactive and reactive aggression.

In their 20-year longitudinal study, Obradović, Burt, and Masten (2010) found evidence for a developmental cascade model in which externalizing behaviors in childhood predicted academic competence in late adolescence, which in turn predicted social competence in emerging adulthood; further, academic and social competence both predicted depressive symptoms from emerging to young adulthood (although the latter pathway held only for males). In a much shorter study, van Lier et al. (2012) followed children from 6 to 8 years of age and found that externalizing problems predicted poor academic performance and peer victimization, both of which in turn predicted internalizing symptoms but with bidirectional associations between externalizing problems and victimization. However, it is difficult to draw firm conclusions from these studies because they investigated similar questions across very different developmental periods and data collection intervals. Moreover, the behavioral, social, academic, and emotional constructs in these studies were measured at a much broader level (e.g., externalizing problems as opposed to reactive aggression) than is suggested by theory.

Interestingly, much of the research reviewed above supports the possibility of alternative mediation sequences that all fall within the framework of the failure model. Specifically, four models are tested in the present study. The original dual pathways model (Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991) shows an assumption that academic performance and peer rejection act as *additive* and *simultaneous* mediators in the developmental pathways from aggression to depressive symptoms. Although most of the research reviewed above supports components of this model, some studies (e.g., van Lier et al., 2012) provide specific support for academic performance and peer rejection as simultaneous, dual-pathway mediators.

However, other studies suggest that these mediators may follow a particular sequence within the developmental pathway rather than occurring simultaneously. In fact, the extant research and theory supports two possible sequences, such that peer rejection could predict poor academic performance or vice versa. In support of the former, Fite and colleagues found that peer rejection accounted for the cross-sectional association between reactive (but not proactive) aggression and academic performance (Fite et al., 2013); and in a separate sample, peer rejection was the only significant longitudinal predictor of academic performance (Fite et al., 2012b). These two studies, taken together, suggest that the direction of the associations are such that reactive aggression predicts peer rejection, which in turn predicts poor academic performance, which ultimately contributes to depressive symptoms. On the other hand, the revised model favored by Patterson and Capaldi (1991) suggests the reverse three-chain mediation sequence, where poor academic performance precedes peer rejection. The developmental cascade identified by Obradović et al. (2010)—from externalizing to academic to social to internalizing problems—provides some support for this model, but with the model unfolding over a later developmental period and a longer timeframe from late childhood through adulthood.

Finally, other longitudinal studies (e.g., Fite et al., 2014; Schwartz et al., 2008) highlight the importance of social functioning as a moderator of longitudinal associations among aggression, academic performance, and depressive symptoms. Most notably, Schwartz et al. (2008) found that, in a sample of school-age children, having fewer friends and a lower GPA were both predictive of depressive symptoms the following year; however, there was a significant interaction between these social and academic variables, such that success in either domain (e.g., having more friends) served to cancel out the adverse effects of the other domain (e.g., a lower GPA) in contributing to depressive symptoms. These findings suggest a fourth

possibility, moderated mediation, where social and academic functioning interact to buffer or exacerbate risk for depressive symptoms. Thus, the nature, sequence, and relative magnitude of these developmental associations remain unclear, underscoring the need for rigorous, longitudinal examinations of alternative models.

It should also be noted that the research reviewed above utilized s a wide array of measurement methodologies, including teacher, parent and self-report, observer impressions, peer nominations, and test scores. For example, Day et al. (1992) relied entirely on teacher report for ratings of proactive/reactive aggression, academic performance, social functioning, whereas Morrow et al. (2008) constructed latent variables for proactive/reactive aggression, depressive symptoms, and peer rejection, each with four indicators representing different informants (parent, teacher, peer, self). More commonly, researchers have used observed scores (e.g., Fite et al., 2012b; Panak & Garber, 1992; Van Lier & Koot, 2010) or latent constructs (e.g., Obradovic et al., 2010; Patterson & Stoolmiller, 1991, Van Lier et al., 2012) comprised of informants/ methods that differed across variables, introducing additional method variance. While latent variable designs offer a number of strengths in terms of reducing measurement error, it can also be useful for inferential and application purposes to test models using data from a particular perspective. The present study contributes to the literature utilizing teacher report data, which is particularly useful given that teachers represent a significant and common perspective for identifying and referring youth for social, behavioral, and academic problems.

Study Overview and Hypotheses

In sum, the literature reviewed above does not disconfirm, but rather provides mixed evidence for the failure model posited by Patterson and Capaldi (1990). As described in the preceding section, the original model (presented conceptually in Figure 1) can be translated into

four variations (presented analytically in Figure 2, Models 1-4), each with some degree of theoretical and empirical support. Specifically, these include (a) the original dual-pathway model, in which peer rejection and poor academic performance are simultaneous, additive predictors of depressive symptoms (Figure 2, Model 1); (b) a social cascade model, in which peer rejection predicts poor academic performance (Figure 2, Model 2); (c) an academic cascade model, in which poor academic performance predicts peer rejection (Figure 2, Model 3); and an interaction model, in which peer rejection and poor academic performance interact to buffer or exacerbate risk for depressive symptoms (Figure 2, Model 4).

The goal of the present study is to provide a direct, longitudinal examination of the failure model, including a theory-driven, comparative evaluation of the four variations described above. A further contribution of this study is to examine distinct developmental pathways associated with the proactive and reactive subtypes of aggression. Specifically, this study investigates these models among a sample of school-age children over the course of three years, within an accelerated design spanning first grade through fifth grade.

Based on previous research, it was hypothesized that: (a) reactive, not proactive, aggression will predict subsequent peer rejection and academic performance; (b) peer rejection and academic performance will both predict subsequent depressive symptoms; and (c) the interaction of peer rejection and academic performance will predict depressive symptoms, such that poor academic performance combined with higher levels of peer rejection will be associated with the highest levels of depressive symptoms. Given the mixed findings regarding the sequence of peer rejection and academic performance as mediators, and the possible interaction between social and academic functioning, no specific hypothesis was made regarding

comparisons of the paths and model fit of alternative models. Rather, the present study is intended to help resolve this issue.

Finally, it is important to consider that the studies reviewed above have used male-only and mixed-gender samples, with results somewhat mixed or unclear with regard to gender-related differences. Accordingly, the present study explores whether the significance or magnitude of the model paths differs between boys and girls. Given the limitations of the available evidence, no specific hypotheses were made regarding gender differences.

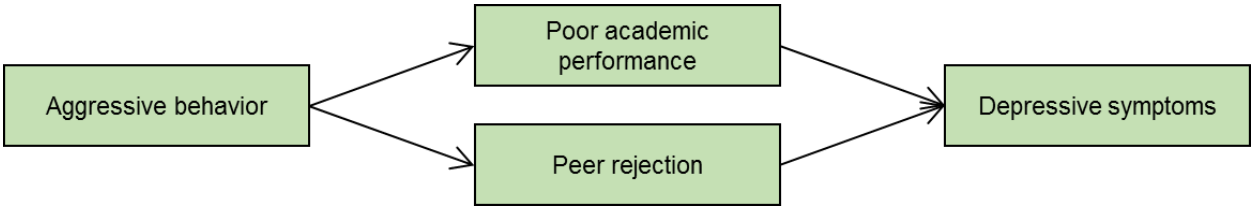


Figure 1. Conceptual diagram of the dual pathway failure model
As originally conceptualized (Patterson & Capalidi, 1990; Patterson & Stoolmiller, 1991), poor academic performance and peer rejection both mediate the association between early aggressive behavior and subsequent depressive symptoms.

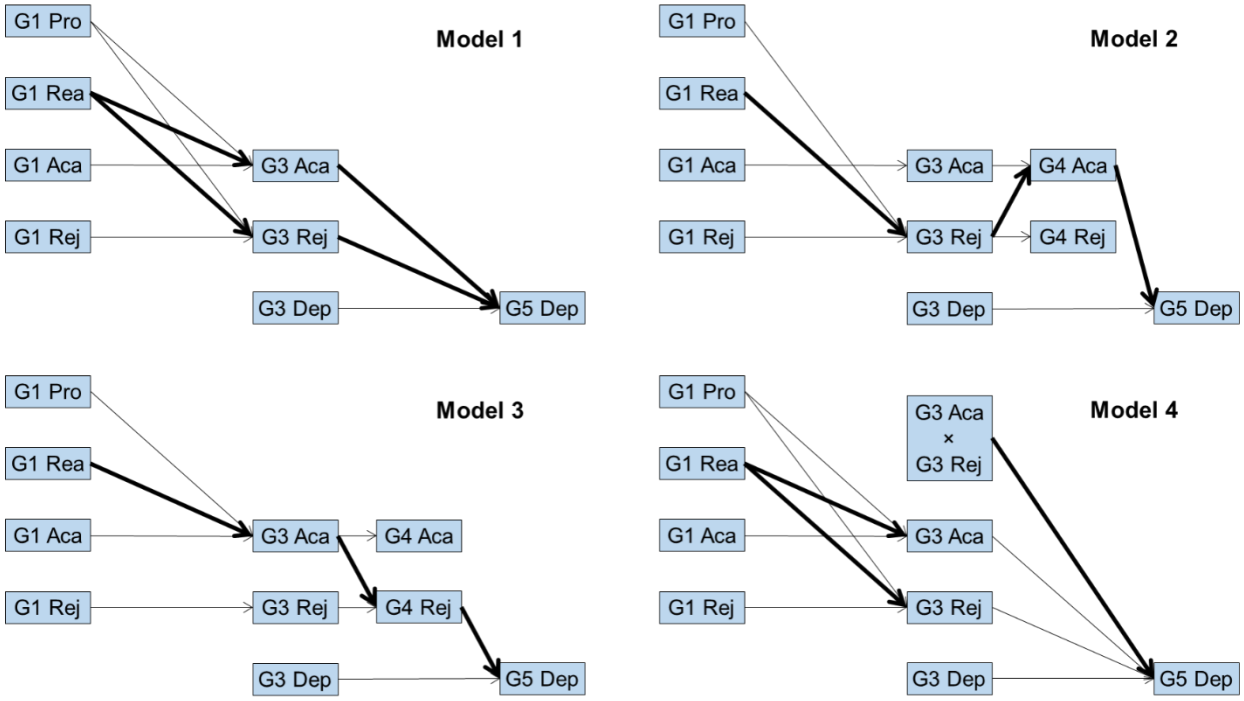


Figure 2. Analytic diagrams of the four models tested in the present study
Bold lines denote paths that would need to be significant in order for model to be fully supported; gray lines depict all other regression paths in the model. All models control for gender but these paths are not depicted for clarity purposes. G1…G5 = 1st grade…5th grade; pro = proactive aggression, rea = reactive aggression, aca = academic performance, rej = peer rejection, dep = depressive symptoms.

Method

Participants

Data collection took place as part of a larger research project in partnership with administrators and teachers at an elementary school in a small city in the U.S. Midwest. Participants were the primary classroom teachers of students in grades 1-5 (ages 6-12), who were asked to complete measures on each student in their class. As described in more detail below, these data were collected annually over the course of three academic school years. In total, data were collected from 36 unique teachers reporting on 963 unique students (48.7% girls) on at least one of the three occasions. According to school records from the 2012-2013 academic year, 88% of the students were White and 38% were eligible for free or reduced lunch.

Longitudinal Design

This study follows an accelerated longitudinal design over the course of three academic school years. Specifically, data were collected between late October and early December of each fall semester in 2012 (T1), 2013 (T2), and 2014 (T3). The present analyses utilize data on students who were in grades 1, 3, 4, or 5 on any of these time points. As part of the accelerated design, the longitudinal metric of change was grade level (1st, 3rd, 4th, 5th) rather than timing of data collection. This design was selected because the focus of the present study is on developmental pathways that are bound to specific developmental periods (i.e., aggression during early elementary school, depressive symptoms during late elementary school, and social and academic functioning in the interim), and this aspect of the study is of much greater interest than the years in which the data collection occurred. Moreover, given the aims and hypotheses, it would be inappropriate to combine a range of grade levels (e.g., K – 3rd), simply because the data were collected at the same time. The accelerated design results in a transformed data

structure, which is illustrated in Table 1 along with child sample sizes and teacher response rates by cohort and grade level. For example, Cohort C students were enrolled in third grade in fall 2012; Cohort D students were in third grade the following year, and Cohort E the year after that. Thus, observations of third grade as the metric of change were aggregated from 2012, 2013, and 2014 data collection time points, respectively, for Cohorts C, D, and E.

The four hypothesized models are presented above in Figure 2. The equal intervals between the annual fall data collections (G1 – G3 – G5) were designed primarily for an examination of the dual pathway models, as equal intervals are strongly advised for longitudinal mediation models (Cole & Maxwell, 2003). Overall, physical aggression decreases from early through late childhood (Vitaro & Brendgen, 2012), whereas levels of depressive symptoms increase as children approach early adolescence (Crijnen, Achenbach, & Verhulst, 1999). Thus, one-year intervals were selected in order to capture these processes during key developmental periods. As shown in Figure 2, peer rejection and academic performance were modeled as occurring within third and fourth grade in Models 2 and 3. These cascade sequences were investigated during these years because this period represents a key period of academic and social development where cascading bidirectional effects may be observed from one year to the next. At the same time, the timing of the measurement for all baseline (G1) and outcome (G5) variables remains consistent across all models. Thus, necessary variations in data collection timing notwithstanding, comparisons between Models 1 and 4 (G1 – G3 – G5) on the one hand, and between Models 2 and 3 (G1 – G3 – G4 – G5) on the other hand, collectively facilitate a rigorous evaluation of four variations of the failure model.

Table 1. Data collection samples and occasions by cohort and grade level

Study Cohort	Grade Level					Total
	1st	2nd	3rd	4th	5th	
A	-	-	-	-	104 [a] (5/5)	104 (5/5)
B	-	-	-	137 [a] (7/7) →	143 [b] (6/6)	158 (13/13)
C	-	-	119 [a] (6/6) →	116 [b] (5/5) →	117 [c] (5/5)	137 (16/16)
D	-	126 [a] (6/6) →	124 [b] (6/6) →	118 [c] (6/6)	-	143 (16/16)
E	111 [a] (5/6) →	139 [b] (7/7) →	140 [c] (6/6)	-	-	161 (18/19)
F	122 [b] (7/7) →	123 [c] (6/6)	-	-	-	132 (13/13)
G	128 [c] (7/8)	-	-	-	-	128 (7/8)
Total	361 (19/21)	388 (19/19)	383 (18/18)	371 (18/18)	364 (16/16)	963 (90/92)

Note: Child sample sizes are reported outside parentheses, with classroom/teacher response rates inside parentheses. Total Grade Level *N* reflects the sum total of the figures within each column, summing to 100%. Total Cohort *N* reflects the total number of unique students within each cohort measured on at least one occasion and therefore do not sum to 100%. Gray shading denotes data that were collected but are not used in the present analyses. [a] Fall 2012 data collection. [b] Fall 2013 data collection. [c] Fall 2014 data collection.

Procedures

The same protocol was used for data collection at each of the three time points. Prior to each wave of data collection, teachers were informed of the nature of their participation and potential benefits to the school and to themselves. As shown in Table 1, virtually all teachers in grades 1-5 provided consent and participated in the study at each time point, yielding a total grade-by-year, classroom-level response rate of 97.8%. The only exceptions to the otherwise 100% response rate were first grade teachers at T1 and T3 (83.3% and 87.5% participation, respectively). Teachers completed a series of measures, hosted through the Qualtrics online survey platform, in reference to each student in their classroom. Teachers were given a list of their students' names and randomly assigned temporary ID numbers for each study. These temporary IDs were used by the teachers to de-identify which student was being rated, and later by the research team to match up with participant's actual study ID numbers. The study ID numbers were then used create the longitudinal dataset, including between 1 and 3 ratings of each student over time. Teachers received $7 for each survey completed at T1 and $50 for their full participation at each subsequent time point.

Importantly, these teachers routinely evaluate and report on their students in these domains (i.e., academic performance, social-emotional functioning, behavior problems) as part of their regular professional duties and school policies and procedures. Further, the data were de-identified to the researchers (or anyone who inadvertently saw the database) and were only presented back to teachers and administrators in aggregate form (i.e., frequencies and averages schoolwide or by grade level). For these reasons, it was determined by the researchers, the institutional review board, and the school administrators that teacher participation was likely to have no effect on individual students, thereby rendering teacher consent sufficient for the data

being collected (and parent consent and youth assent unnecessary). This study was approved by the Human Subjects Committee – Lawrence (HSCL) at the University of Kansas (HSCL #20175).

Measures[1]

Demographic information. Teachers were asked to report the gender and grade level of each student in their class.

Proactive and reactive aggression. The Proactive and Reactive Aggression rating scale (PRA; Dodge & Coie, 1987) was used to collect teachers' ratings of student's aggressive behaviors. The PRA consists of six items, with subscales measuring proactive (e.g., threatens or bullies others to get what s/he wants) and reactive aggression (e.g., when teased or threatened, s/he gets angry easily and fights back). All items are rated on a five-point Likert scale from 1 (*never*) to 5 (*almost always*). Proactive and reactive aggression scores were calculated by taking the averages of the items within the respective subscales. Past research has supported the validity and reliability of the PRA as a teacher-rated measure of aggressive behavior in school-age children (Dodge & Coie, 1987; Dodge, Lochman, Harnish, Bates, & Pettit, 1997). Across all time points, internal consistency was at good for proactive aggression (αs = .85-.87) and excellent for reactive aggression (αs = .93-.95).

Academic performance. Academic performance was measured using items rated by the students' primary classroom teacher. Items asked about the child's academic performance in three ways: (a) "relative to other students in your class," (b) "overall academic performance (reputation based on all their classes)," and (c) "what letter grade best reflects this student's academic performance." All three items were rated on five-point Likert scales, based on the

[1] All measures are included in Appendix A.

child's performance in comparison to peers for the first two items (1 = *well below average*, 2 = *below average*, 3 = *average*, 4 = *above average*, 5 = *well above average*), and using letter grades for the third item (A, B, C, D, or F; reverse-coded as 1-5 for consistency with the other two items). Scores were calculated as the average of these three items, such that higher scores reflect better overall academic performance. Similar items for academic performance have demonstrated evidence of reliability and validity in previous research (e.g., Becker et al., 2014; Evans et al., 2016; Fite et al., 2013). These items demonstrated excellent internal consistency (αs = .93-.95).

Peer rejection. Peer rejection was measured using teachers' ratings on four items from the Teacher Report Form (TRF; Achenbach & Rescorla, 2001). Ratings were provided on a three-point Likert scale (1 = *not true*, 2 = *somewhat or sometimes true*, 3 = *very or often true*) and then averaged for analyses. The peer rejection subscale includes four items that tap general aspects of poor peer relations and social difficulties, including being teased, left out, or not liked. Previous research supports the validity of this subscale in terms of convergent, divergent, and criterion-related validity with other measures of social functioning (e.g., Evans et al., 2015, 2016; Fite et al., 2012b, 2013). The peer rejection scale demonstrated acceptable or good internal consistency at each time point (αs = .78-.83).

Depressive symptoms. The withdrawn/depressed subscale from the TRF was used to measure students' depressive symptoms. This scale consists of eight items, rated on the same three-point Likert scale as peer rejection, measuring symptoms of depression (e.g., sadness or depressed mood, anhedonia, psychomotor retardation) as well as behaviors indicating social withdrawal (e.g., preference for being alone and for not talking to others). The withdrawn/depressed subscale is one of the original composite scales in the Achenbach

assessment system, with substantial empirical support for validity and reliability (Achenbach & Rescorla, 2001; Crijnen, Achenbach, & Verhulst, 1999). This measure demonstrated good internal consistency (αs = .87-88).

Analytic Plan

Analyses included all children for whom data were collected during at least one of the four grade levels. With any accelerated design, missing data can be partitioned into two types: (a) *planned missngness*, for example, where first through third grade observations of Cohort B are "missing," but only because these children were in fourth grade at T1 when the study began; and (b) *unplanned missingness*, the more conventional type of missing data, in which an observation was supposed to be collected but for some reason was not (Little, 2013). Given the high rates of participation and the use of forced-response prompts in the online survey, the overall rates of unplanned missing data were relatively low across all time points. Of the students for whom any data were collected at any time point (N = 963), rates of unplanned missingness for study variables by grade level ranged from a minimum of 8.8% missing (fifth grade depressive symptoms) to a maximum of 15.3% (fourth grade peer rejection). There was no missing data for gender, classroom teacher ID, or cohort. The most common reason for unplanned missing data was a teacher's nonparticipation during one year, leading to occasion-specific missing data for students who had been rated by other teachers in other years. This was considered a missing-at-random mechanism (Enders, 2010) because students are assumed to be randomly assigned to classrooms from one year to the next, and one teacher's decision to not participate in the study during one year, while systematic in nature (and therefore not missing completely at random), is most likely not related to characteristics of individual students for whom data are missing.

With respect to planned missingness, consider that a hypothetical version of this study, constructed as a fully prospective cohort-sequential design, would include seven cohorts, each followed across five years (1st to 5th grade), resulting in 35 cohort-by-grade sets of observations (i.e., all of the cells in Table 1). Of these, the present accelerated design includes 15 cohort-by-grade sets of observations, amounting to approximately 57.1% planned missingness. Such rates of planned missingness are not unusual among accelerated longitudinal designs (Little, 2013), and research has demonstrated the utility of accelerated designs and their comparability with true longitudinal designs (Duncan, Duncan, & Hops, 1996). Indeed, whereas the present analyses are based on three years of data, the hypothetical "complete" longitudinal version of this study—that is, tracking the same seven cohorts of children throughout their entire elementary school careers—would require 11 years of data collection.

In the present analyses, missing data (planned and unplanned) were accommodated using full information maximum likelihood estimation (FIML), a model-based approach that yields estimates that are more efficient and less biased than alternative techniques for handling missing data (Enders & Bandalos, 2001). The equivalence or superiority of FIML over ad hoc methods has been generally demonstrated across different types and levels of missingness, even as rates of missing data reach 50% and 75%. (Newman, 2003). However, it should be noted that when large percentages of the data are missing (e.g., 75%), parameter bias and standard error estimates can become unacceptable under all major missing data techniques, including FIML (Newman, 2003). At higher levels of missing data, the inflation of standard errors can contribute to an increased likelihood of a Type II error. Thus, the results of the present accelerated-design study should be interpreted with caution, as preliminary findings warranting replication within a full prospective cohort design.

It was expected that some variables (e.g., proactive aggression, depressive symptoms) would exhibit a positive skew and kurtosis that departed from normality. This was addressed through the use of robust maximum likelihood estimation, which takes non-normality into account when estimating standard errors and chi square tests (Muthén & Muthén, 2012). All variables were mean-centered to reduce multicollinearity between product terms and their constituent variables in moderator analyses (Kline, 2011) and to aid in the interpretation of the unstandardized regression coefficients.

Prior to estimating path models, data were examined for univariate characteristics (means, standard deviations, internal consistency, and normality), bivariate correlations, and missing data patterns for all variables across all time points. Additionally, the possible effects of students being cross-classified within teachers/classrooms were explored (see Appendix B). [2]

True mediation is established when variable X at Time T is associated with variable Y at Time T+2, and both are associated with variable M at time T+1 (Cole & Maxwell, 2003; Little, 2013; Selig & Preacher, 2009). Similarly, three-chain mediation is established when these same mediational sequences are significant within a sequence involving three paths and two mediators ($X \rightarrow M_1 \rightarrow M_2 \rightarrow Y$) across four time points. However, given the accelerated design, the direct paths from first to fifth grade were not estimable because no participants were observed at both time points. Thus, only the indirect paths were estimated, and direct effects could not be examined. The interaction path in Model 4 was identified using the product term of third grade

[2] Classroom was considered but not included as a covariate or multilevel factor for several reasons. Students were nested within up to three classrooms over the course of the study as they transitioned from one grade to the next. This pattern of nesting represents a cross-classification design rather than a hierarchical or multilevel design (Hox, 2002). For the purposes of this study, cross-classification was not of substantive theoretical interest and was not modeled analytically given the already high degree of complexity in the models. Rather, cross-classification was addressed via (a) supplementary analyses of classroom effects, which were generally minimal (see Appendix B); (b) an assumption of random assignment from one year's classroom to the next; (c) aggregation of data across cohorts and grade levels; and (d) careful interpretation of different grade levels and intervals as not directly comparable.

peer rejection and academic performance as an additional predictor of fifth grade depressive symptoms, thereby allowing one variable to moderate the M→Y path for the other variable, and vice versa (Preacher, Rucker, & Hayes, 2007).

Structural equation modeling (SEM) path analyses were conducted using MPlus 7.2 statistical software (Muthén & Muthén, 2012). The four hypothesized models were examined through a theory-driven, model building and model comparison approach (Kline, 2011). That is, minimalist versions of these models were estimated first, including only the essential indirect mediation paths (as depicted in Figure 2), with gender included as a covariate for all variables at each time point. The four models were first examined separately, then comparatively, through considerations of theory, path coefficients, and model fit. With regard to theory and path coefficients, regression estimates were inspected for significance ($\alpha = .05$) to interpret the hypothesized paths in accordance with the criteria described above. The magnitude of path estimates and of the R^2 values for endogenous variables were also examined in order to inform the interpretation of each model's results.

Next, the models were re-examined to identify any differences related to gender using the product-of-coefficients strategy for assessing conditional indirect effects (Preacher, Rucker & Hayes, 2007). That is, gender was included as a moderator of all regression paths in the model by adding its product term (gender × predictor variable) as an additional predictor variable. Any significant moderation paths were interpreted as an indication that the strength of the association between variables is conditional upon the status of gender. These effects were further explored by re-estimating the models as multiple group models, with the gender-moderated paths allowed to vary between boys and girls.

The overall fit of the final models was evaluated based on model test statistics and absolute, comparative, predictive fit indices between models. Specifically, comparisons were first made in pairs (Models 1 & 4 and Models 2 & 3) that share greater the greatest amount of similarity in their path structures (dual pathway models and cascade sequence models, respectively) and data collection occasions (G1 – G3 – G5 and G1 – G3 – G4 – G5, respectively), thus rendering them more directly comparable[3] to one another than to the other models. Following Kline's (2011) recommended procedure for comparing similar but nonhierarchical models drawn from the same data, model fit was evaluated using the Akaike Information Criterion (AIC). The AIC is a parsimony-adjusted predictive fit index, which combines calculations of model estimation and model fit applied to hypothetical replications of the same size and randomly selected from the same population. In comparisons between two models, a smaller (most negative) AIC indicates a better balance of model fit and parsimony, and would therefore be retained as the preferred model (Kline, 2011). Further evaluation of model fit was based on collective consideration of the χ^2 test statistic, Root Mean Square Error of Approximation (RMSEA), and Confirmatory Fit Index (CFI). A model is considered to have acceptable fit when $\chi^2/df \leq 3.0$, CFI $\geq .90$, and RMSEA $\leq .08$ (Hu & Bentler, 1999; Kline, 2011; Little, 2013).

Using the conservative parameters recommended by MacCallum, Browne, and Sugawar (1996) ($\alpha = .05$, $\beta = .80$; $RMSEA_{H0} = .05$, $RMSEA_{H1} = .08$), power analyses indicate a required minimum sample size of 449 for a test of exact fit in the model with the fewest degrees of freedom (Model 1). This sample size also provides sufficient power for detecting hypothesized

[3] Despite this within-model-pair similarity, no two of these four models can be considered nested or hierarchical; that is, one model cannot be reproduced from another model simply by adding paths (model building approach) *or* removing paths (model pruning approach). Therefore, a traditional chi-square difference test would be inappropriate.

regression paths in the most complex model (Models 2 and 3). Thus, even when taking the planned and unplanned missingness into account, the study sample size of 963 was sufficient to investigate study hypotheses.

Results

Descriptive Statistics and Correlations

Univariate characteristics of study variables from first through fifth grade are presented in Table 2 and Figure 3. As shown in Table 2, the distributional characteristics of study variables were largely similar across grade levels. As expected, departures from normality were observed in some study variables, particularly for proactive aggression and peer rejection. This was handled by robust estimation, as noted above. The overall and cohort means of all variables from first through fifth grade are presented visually in Figure 1. Generally, average levels of each variable appear to be similar across cohorts both concurrently and over time. Supplementary analyses confirmed that there were little to no significant differences or group dependencies between cohorts (although there were some small between-group effects of teachers/classroom for some variables in first and third grade; see Appendix B).

With regard to bivariate correlations (Table 3), proactive and reactive aggression were highly correlated with one another (sharing between 45 and 62% of their concurrent variance across grade levels), whereas academic performance and peer rejection showed a small but significant negative correlation (sharing between 3 and 9% of their concurrent variance across grade levels). Notably, ratings of academic performance showed consistently high stability over time (1- and 2-year autocorrelations ranging from .66 to .73), whereas peer rejection (1- and 2-year autocorrelations from .11 to .42) and depressive symptoms (1- and 2-year autocorrelations from .19 to .52) showed relatively lower stability over time and greater variability.

Examining the hypothesized model paths, reactive aggression in first grade was moderately correlated with peer rejection but not academic performance in third grade. First grade proactive aggression showed the same pattern of associations, but with a smaller effect. In

the later elementary school years, third and fourth grade peer rejection were both weakly correlated with fifth grade depressive symptoms. However, neither third nor fourth grade academic performance was significantly associated with fifth grade depressive symptoms. With regard to gender differences in the model paths (Table 2), girls showed lower levels of reactive aggression in first grade, and lower levels of peer rejection in both first and fourth grade. Interestingly, proactive aggression, academic performance, and depressive symptoms were uncorrelated with gender at all grade levels included in the path models.

Table 2. Univariate characteristics of study variables at each grade level

Variable	*N*	*M*	*SD*	Observed Range	Theoretical Range	Skewness	Kurtosis	*r* with Female
G1 Proactive Aggression	361	1.38	0.76	1.00-4.67	1-5	2.25	4.53	-.09
G2 Proactive Aggression	388	1.20	0.55	1.00-4.00	1-5	3.26	10.50	-.13*
G3 Proactive Aggression	383	1.19	0.52	1.00-4.67	1-5	3.46	13.70	-.03
G4 Proactive Aggression	371	1.16	0.46	1.00-4.33	1-5	3.71	15.15	-.11*
G5 Proactive Aggression	364	1.30	0.56	1.00-3.67	1-5	2.00	3.59	-.09
G1 Reactive Aggression	361	2.07	1.21	1.00-5.00	1-5	0.94	-0.20	-.22**
G2 Reactive Aggression	388	1.61	1.02	1.00-5.00	1-5	1.76	2.13	-.15**
G3 Reactive Aggression	383	1.45	0.91	1.00-5.00	1-5	2.26	4.56	-.19**
G4 Reactive Aggression	371	1.52	0.93	1.00-5.00	1-5	2.06	3.64	-.25**
G5 Reactive Aggression	364	1.56	0.87	1.00-5.00	1-5	1.78	3.03	-.17**
G1 Peer Rejection	361	1.17	0.32	1.00-2.75	1-3	2.31	5.27	-.13*
G2 Peer Rejection	388	1.13	0.30	1.00-3.00	1-3	3.08	10.83	-.08
G3 Peer Rejection	382	1.16	0.34	1.00-3.00	1-3	2.63	7.33	.07
G4 Peer Rejection	371	1.13	0.32	1.00-3.00	1-3	2.92	8.85	-.13*
G5 Peer Rejection	364	1.19	0.37	1.00-2.75	1-3	2.26	4.82	-.06
G1 Academic Performance	361	3.18	1.18	1.00-5.00	1-5	-0.10	-0.98	.00
G2 Academic Performance	388	3.44	1.07	1.00-5.00	1-5	-0.53	-0.60	.03
G3 Academic Performance	381	3.50	1.09	1.00-5.00	1-5	-0.34	-0.54	.07
G4 Academic Performance	372	3.55	1.03	1.00-5.00	1-5	-0.53	-0.42	.09
G5 Academic Performance	363	3.68	0.94	1.33-5.00	1-5	-0.41	-0.47	.10
G1 Depressive Symptoms	361	1.19	0.31	1.00-2.75	1-3	2.27	5.31	-.02
G2 Depressive Symptoms	388	1.24	0.39	1.00-3.00	1-3	2.04	3.73	-.01
G3 Depressive Symptoms	382	1.23	0.37	1.00-3.00	1-3	2.03	4.13	-.02
G4 Depressive Symptoms	371	1.17	0.32	1.00-2.75	1-3	2.21	4.49	-.06
G5 Depressive Symptoms	364	1.17	0.29	1.00-2.38	1-3	2.14	4.28	-.04

Note: Gray shading denotes variable not used in models, presented here for descriptive purposes. *$p < .05$, **$p < .01$

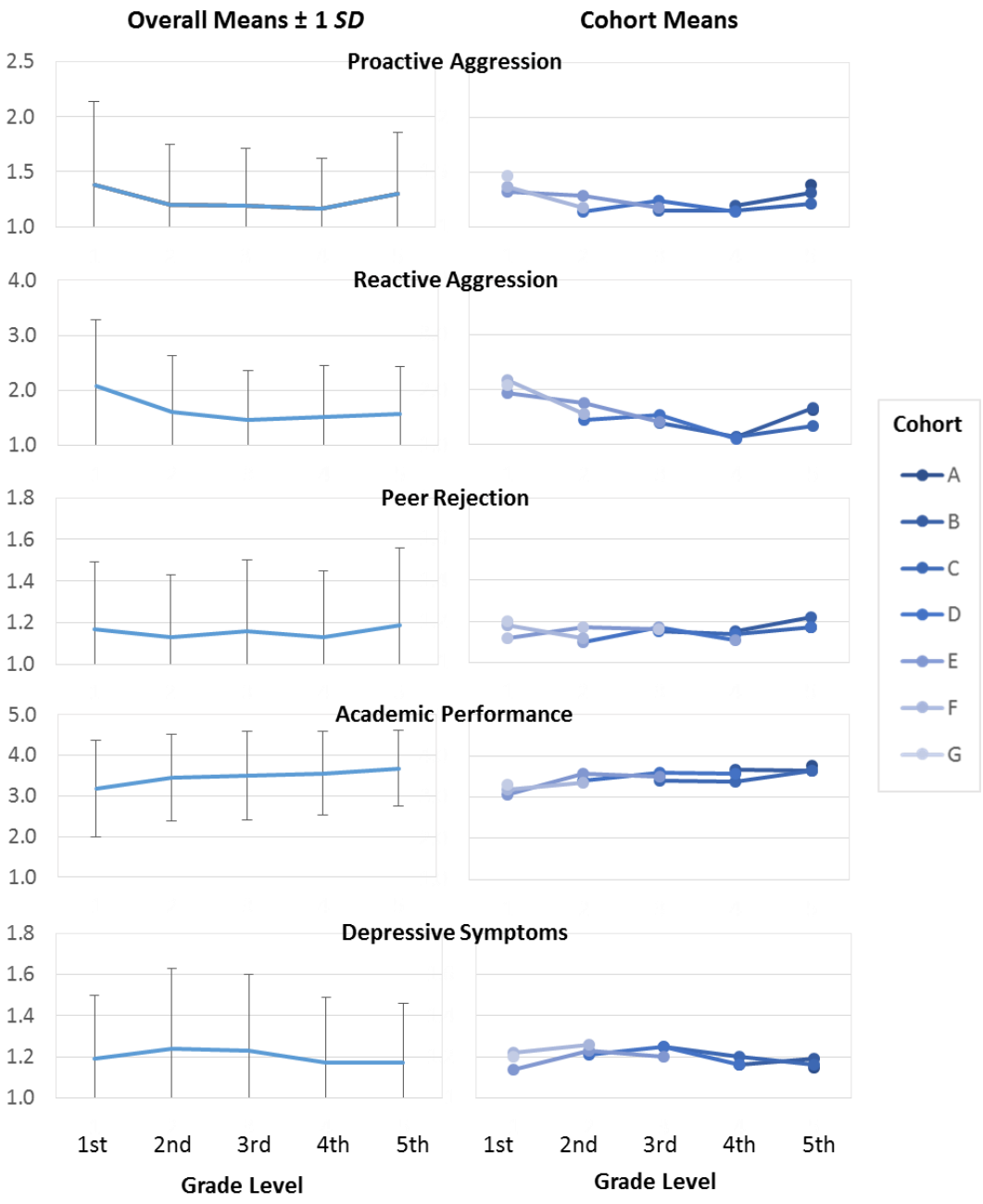


Figure 3. Average scores on study variables across grade levels, overall and by cohort

Table 3. Bivariate correlations of all study variables across all five grade levels

		First Grade					Second Grade					Third Grade					…
		Pro	Rea	Aca	Rej	Dep	Pro	Rea	Aca	Rej	Dep	Pro	Rea	Aca	Rej	Dep	…
G1	Pro	–															
	Rea	.74**	–														
	Aca	-.16**	-.22**	–													
	Rej	.55**	.63**	-.18**	–												
	Dep	.09	.07	-.15**	.28**	–											
						(*n* = 361)											
G2	Pro	.21**	.21**	.02	.14*	-.07	–										
	Rea	.20**	.27**	-.13	.18**	-.06	.79**	–									
	Aca	-.09	-.25**	.72**	-.21**	-.04	-.03	-.10*	–								
	Rej	.06	.12	-.20**	.11	-.10	.61**	.66**	-.19**	–							
	Dep	-.04	-.09	-.28**	-.05	.20**	.16**	.20**	-.29**	.35**	–						
						(*n* = 213)					(*n* = 388)						
G3	Pro	.14	.17	.03	.15	.01	.30**	.31**	-.07	.34**	.00	–					
	Rea	.23*	.35**	-.04	.30**	-.01	.27**	.33**	-.11	.33**	.04	.77**	–				
	Aca	-.06	-.16	.66**	-.11	.00	-.09	-.13*	.71**	-.24**	-.32**	-.09	-.15**	–			
	Rej	.21*	.31**	-.05	.21*	-.03	.35**	.35**	-.12	.42**	.09	.54**	.67**	-.21**	–		
	Dep	.14	.31**	-.24*	.33**	.24*	.04	.11	-.16*	.11	.28**	.07	.23**	-.26**	.45**	–	
						(*n* = 96)					(*n* = 243)					(*n* = 383)	
G4	Pro	–	–	–	–	–	.36**	.34**	-.19	.42**	.13	.34**	.32**	-.06	.21**	-.05	…
	Rea	–	–	–	–	–	.38**	.46**	-.16	.45**	.04	.41**	.59**	-.11	.39**	.05	…
	Aca	–	–	–	–	–	-.24*	-.20*	.72**	-.15	-.37**	-.15*	-.20**	.72**	-.19**	-.09	…
	Rej	–	–	–	–	–	.27**	.33**	-.06	.34**	.02	.32**	.45**	-.12	.39**	.12	…
	Dep	–	–	–	–	–	.02	.10	-.01	.08	.19	.05	.16*	-.11	.27**	.36**	…
						(*n* = 0)					(*n* = 102)					(*n* = 220)	
G5	Pro	–	–	–	–	–	–	–	–	–	–	.30**	.22*	-.14	-.05	-.17	…
	Rea	–	–	–	–	–	–	–	–	–	–	.28**	.43**	-.07	.09	-.13	…
	Aca	–	–	–	–	–	–	–	–	–	–	-.20*	-.22*	.73**	-.09	-.10	…
	Rej	–	–	–	–	–	–	–	–	–	–	.16	.40**	-.13	.36**	.09	…
	Dep	–	–	–	–	–	–	–	–	–	–	-.02	.05	-.03	.27**	.38**	…
						(*n* = 0)					(*n* = 0)					(*n* = 100)	

		Fourth Grade					Fifth Grade				
		Pro	Rea	Aca	Rej	Dep	Pro	Rea	Aca	Rej	Dep
G4	Pro	–									
	Rea	.67**	–								
	Aca	-.18**	-.29**	–							
	Rej	.52**	.72**	-.30**	–						
	Dep	.07	.16**	-.17**	.30**	–					
						(*n* = 371)					
G5	Pro	.48**	.42**	-.15*	.25**	.07	–				
	Rea	.39**	.57**	-.18**	.43**	.10	.76**	–			
	Aca	-.12	-.25**	.73**	-.23**	-.12	-.22**	-.28**	–		
	Rej	.25**	.49**	-.16*	.40**	.22**	.50**	.72**	-.25**	–	
	Dep	-.08	.03	-.08	.14*	.52**	.06	.16**	-.11*	.31**	–
						(*n* = 228)					(*n* = 364)

Note: Correlations are based on observed data only using pairwise deletion. Due to the planned missingness of the accelerated design, sample size and membership varies between grade levels. Sample sizes are indicated for all variables within and between grade levels. Correlations spanning more than three grade levels (1st to 4th, 1st to 5th, 2nd to 5th) are not estimated because there were no longitudinal observations that spanned these periods. Gray shading denotes variable not used in models, presented here for descriptive purposes. *$p < .05$, **$p < .01$

Model Estimation

Results of the four hypothesized path models are presented in Figure 4, with model statistics presented in Table 4. As shown in Figure 4, Panel A, results of Model 1 indicated that reactive aggression in first grade predicted lower academic performance and higher peer rejection in third grade. Conversely, first grade proactive aggression positively predicted third grade academic performance, and showed a marginally statistically significant negative association with third grade peer rejection. With regard to the later paths, neither academic performance nor peer rejection in third grade showed any unique association with depressive symptoms in fifth grade. The stability paths were significant for depressive symptoms (G3 → G5) and academic performance (G1 → G3), with a marginally statistically significant path for peer rejection (G1 → G3). Model 2 (Figure 4, Panel B) revealed a significant positive association between reactive aggression in first grade and peer rejection in third grade, whereas proactive aggression showed a marginally statistically significant negative association with subsequent peer rejection. The only other significant paths in Model 2 were stability paths, which were all significant for academic performance (G1 → G3 → G4), peer rejection (G1 → G3 → G4), and depressive symptoms (G3 → G5).

Results of Model 3 (Figure 4, Panel C) show two marginal trends for academic performance in third grade being positively predicted by proactive aggression and negatively predicted by reactive aggression in first grade. As in Model 2, all of the stability paths in Model 3 were significant; these were the only paths that reached statistical significance in this model. Finally, Figure 4 Panel D presents the results of Model 4. As in Model 1, reactive aggression in first grade predicted lower academic performance and higher peer rejection in third grade, whereas proactive aggression in first grade predicted higher academic performance in third grade

and showed a marginal trend for a negative association with peer rejection in third grade. The cross-product term between G3 academic performance and G3 peer rejection did not emerge as significant. Thus, results fail to reject the null hypothesis regarding the possibility of an interaction between academic performance and peer rejection in predicting subsequent depressive symptoms.

Overall, these four models were similar with regard to the amount of variance explained in third, fourth, and fifth grade outcome variables (see R^2s in Figure 4). Although not depicted in Figure 4, gender was included as a covariate of the outcome variables (i.e., modeled as a time-invariant predictor of G3/G4 academic performance, G3/G4 peer rejection, and G5 depressive symptoms) in all models, but showed no significant unique associations in any of those models (*p*s = .11 - .85). Across all models, the substantive paths from reactive aggression in first grade to peer rejection in third grade remained significant, and the stability paths for academic performance and depressive symptoms were also robust across models. By contrast, across models, regression estimates fluctuated between statistical significance ($p < .05$) and marginal statistical significance ($.05 < p < .10$) for the paths from proactive aggression in first grade to both academic performance and peer rejection in third grade, as well as for the paths from first grade reactive aggression to third grade academic performance. The stability paths for peer rejection were robust between third and fourth grade, but showed mixed results across models from first to third grade. These discrepant path results across models make some sense given that the regression equations for any outcome (e.g., G3 peer rejection) include different predictors depending on what model one is viewing (e.g., G3 peer rejection has four predictors in Models 1, 2, and 4, but only two predictors in Model 3), as well as the differences in the samples due as a result of the accelerated study design (Models 1 and 4, $N = 817$; Models 2 and 3, $N = 782$).

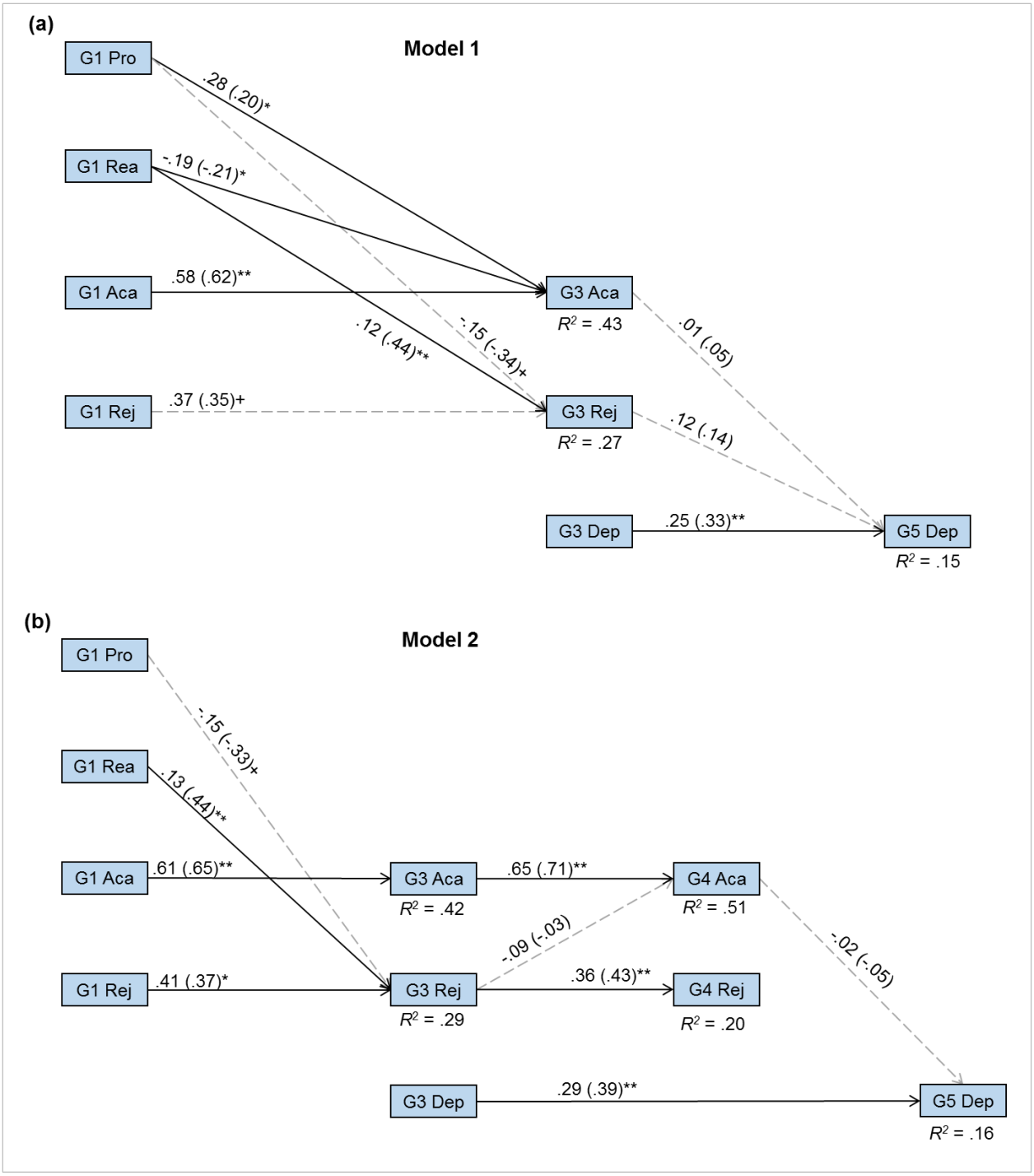


Figure 4. Results of four alternative path models

(Continues on next page.)

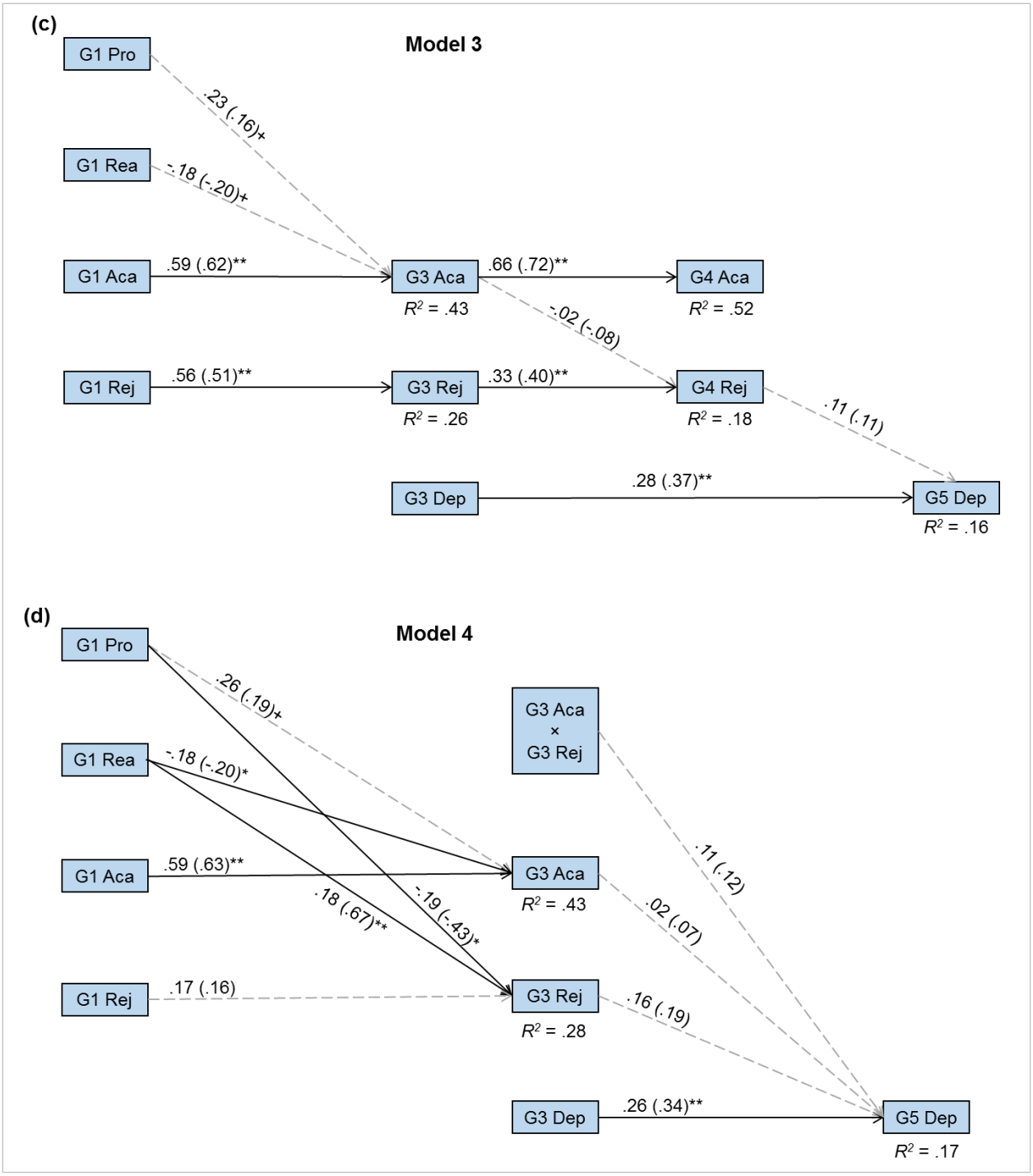


Figure 4. Results of four alternative path models

Black lines denote paths that are significant at $p < .05$; gray dashed lines denote nonsignificant paths ($p \geq .05$). Unstandardized path coefficients are reported outside parentheses, with standardized estimates inside parentheses. Gender was included as a covariate for all endogenous variables; thus, any gender effects are therefore included in the R^2 estimates. + $p < .1$, * $p < .05$, ** $p < .01$.

Table 4. Fit statistics for all models tested

Model	N	N Free Parameters	AIC	BIC	SABIC	χ^2	df	RMSEA	RMSEA 95% C.I.	CFI	TLI	SRMR
Original Models												
1	817	18	-2702.2	-2617.5	-2674.7	52.08	5	.107	.080 - .135	.689	-.059	.056
2	782	24	-1594.8	-1482.9	-1559.1	75.21	14	.075	.059 - .092	.808	.617	.068
3	782	24	-1590.7	-1478.8	-1555.0	73.45	14	.074	.058 - .091	.814	.628	.071
4	817	19	-2939.8	-2850.4	-2910.8	70.84	7	.106	.084 - .129	.644	-.019	.059
With Gender as Moderator												
1b	817	25	-2565.7	-2448.1	-2527.4	60.71	15	.086	.064 - .110	.727	.490	.100
2b	782	33	-1479.1	-1325.3	-1430.1	104.36	33	.074	.059 - .091	.794	.713	.108
3b	782	30	-1460.4	-1320.6	-1415.9	124.84	36	.079	.065 - .095	.744	.672	.125
4b	817	26	-2748.9	-2626.5	-2709.1	86.75	20	.090	.071 - .110	.634	.377	.092
Final Model												
5a	782	40	-1483.2	-1296.7	-1423.8	83.74	26	.075	.058 - .094	.833	.705	.097
5b	782	36	-1489.6	-1321.8	-1436.1	84.27	30	.068	.051 - .085	.843	.760	.106

Gender as Moderator

In light of the varied results described above, models were assessed for gender differences prior to any further evaluation or modifications. Indeed, each model revealed between 1 and 4 paths that were moderated by gender. Results of gender interaction analyses are presented in Table 5.

In Model 1, gender moderation was found for the following paths: third grade depressive symptoms to fifth grade depressive symptoms (B = .45, p = .010), and first grade proactive aggression to third grade peer rejection (B = -.58, p = .004). Additionally, marginally statistically significant interactions were found for third grade peer rejection to fifth grade depressive symptoms (B = -.30, p = .083) and from first grade proactive aggression to third grade academic performance (B = .64, p = .066). In Model 2, significant gender interactions were found for the following paths: third grade depressive symptoms to fifth grade depressive symptoms (B = .31, p

= .037), third grade peer rejection to fourth grade academic performance (*B* = .62, *p* = .010), and first grade proactive aggression to third grade peer rejection (*B* = -.60, *p* = .009).

In Model 3, only one marginally statistically significant gender interaction was found: third grade depressive symptoms to fifth grade depressive symptoms (*B* = .24, *p* = .053). Lastly, in Model 4, the following paths were found to be significantly moderated by gender: third grade depressive symptoms to fifth grade depressive symptoms (*B* = .43, *p* = .016), first grade proactive aggression to third grade peer rejection (*B* = -.59, *p* = .002). A marginally statistically significant interaction was found from first grade reactive aggression to third grade peer rejection (*B* = .21, *p* = .076).

In sum, when looking across models, there were pronounced and robust interactions with gender in the model paths predicting peer rejection in third grade (largely due to first grade proactive aggression, with a possible effect of reactive aggression), and in the model paths leading to depressive symptoms in fifth grade (largely due to third grade depressive symptoms, with a possible effect of peer rejection). There was also a significant interaction in the path from peer rejection in third grade to academic performance in fourth grade, although this was path was only specified for Model 2.

These five interactions were further evaluated by re-estimating the models as multiple-group models with the gender-moderated paths identified above allowed to differ between boys and girls.[4] Given that some regression paths oscillated between statistical significance and marginal trends depending on the structure of the model, gender differences were explored by making the same modifications across all four models, when justified by at least one model (i.e.,

[4] The possible gender × proactive aggression interaction in predicting third grade academic performance (Model 1, *p* = .066) was not considered further given its marginal significance and the absence of any other conditional effects of gender on predictors of third grade academic performance.

the paths leading to third grade peer rejection and to fifth grade depressive symptoms were allowed to vary across genders in all four models, given evidence of moderation by gender in at least one model).

Results of these revised models are presented in Figure 5. In all instances of gender-moderated paths, results were indeed substantively different between boys and girls. For girls, higher levels of proactive aggression in first grade predicted lower levels of peer rejection in third grade; however, no significant effect was found for boys (revised models 1, 2, and 4). Further, for both boys and girls, reactive aggression in first grade predicted peer rejection in third grade; but this effect appears to be stronger for girls (revised models 1, 2, and 4). Interestingly, third grade peer rejection predicted significantly *poorer* fourth grade academic performance among boys, but marginally *better* academic performance among girls (revised model 2 only). Lastly, in boys, peer rejection, but not depressive symptoms, in third grade predicted depressive symptoms in fifth grade; conversely, for girls, depressive symptoms, but not peer rejection, in third grade predicted depressive symptoms in fifth grade (revised models 1, 3, and 4).

Table 5. Examining gender as moderator of all model paths (p-values)

Model Path		Model			
Predictor	Outcome	Model 1	Model 2	Model 3	Model 4
G1 Pro	G3 Aca	.066	-	>.1	>.1
G1 Rea	G3 Aca	>.1	-	>.1	>.1
G1 Aca	G3 Aca	>.1	>.1	>.1	>.1
G1 Pro	G3 Rej	.004	.009	-	.002
G1 Rea	G3 Rej	>.1	>.1	-	.076
G1 Rej	G3 Rej	>.1	>.1	>.1	>.1
G3 Aca	G4 Rej	-	-	>.1	-
G3 Rej	G4 Rej	-	>.1	>.1	-
G3 Rej	G4 Aca	-	.010	-	-
G3 Aca	G4 Aca	-	>.1	>.1	-
G3 Aca	G5 Dep	>.1	-	-	>.1
G3 Rej	G5 Dep	.083	-	-	>.1
G3 Dep	G5 Dep	.010	.037	.053	.016
G3 Aca × Rej	G5 Dep	-	-	-	>.1
G4 Aca	G5 Dep	>.1	>.1	-	>.1
G4 Rej	G5 Dep	>.1	-	>.1	>.1

Note: P-values are reported for the regression coefficient for the interaction term (gender × predictor) all added to the original models simultaneously. Specific values are reported only when statistically significant ($p < .05$) or marginally statistically significant p-values ($p < .1$).

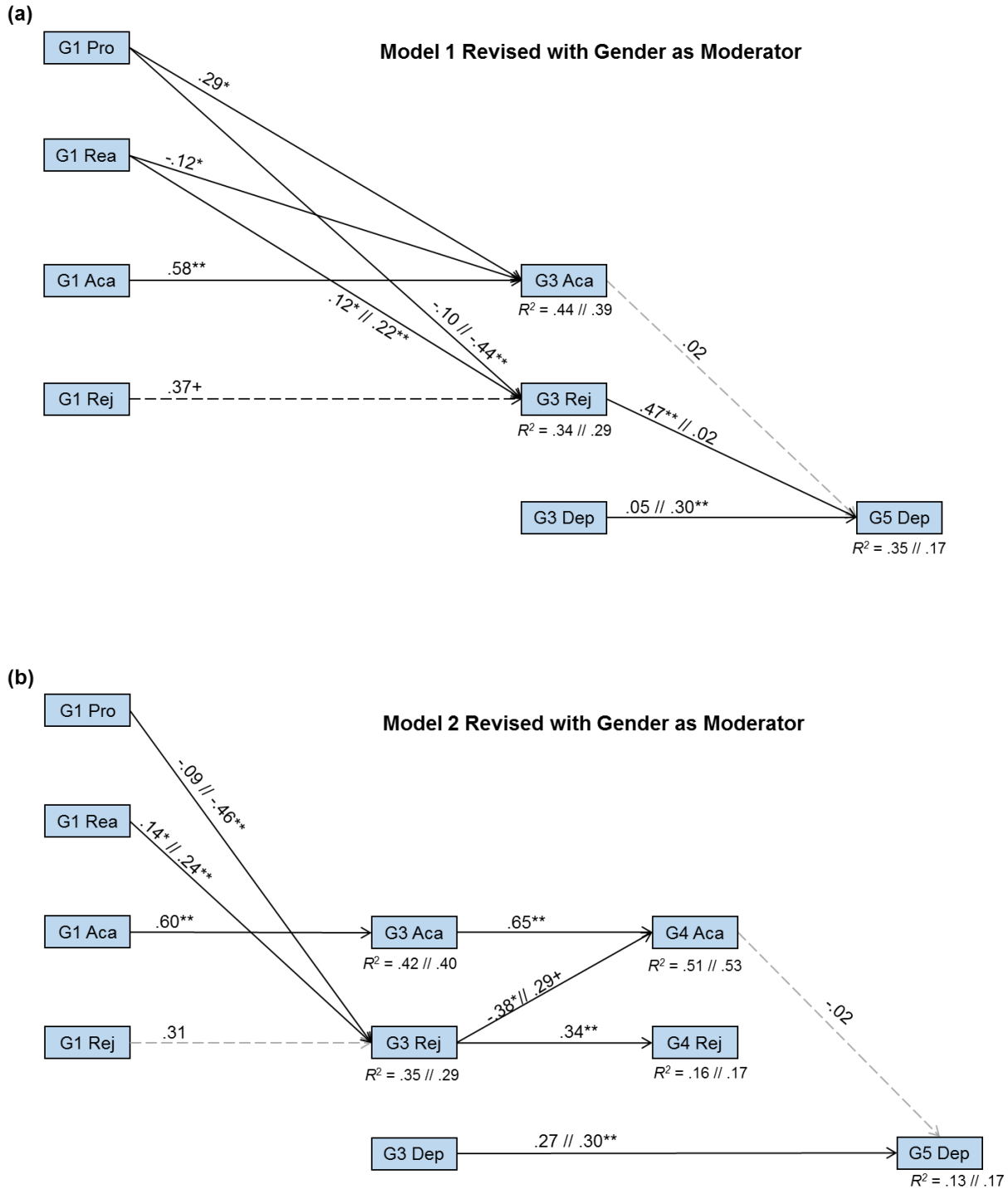
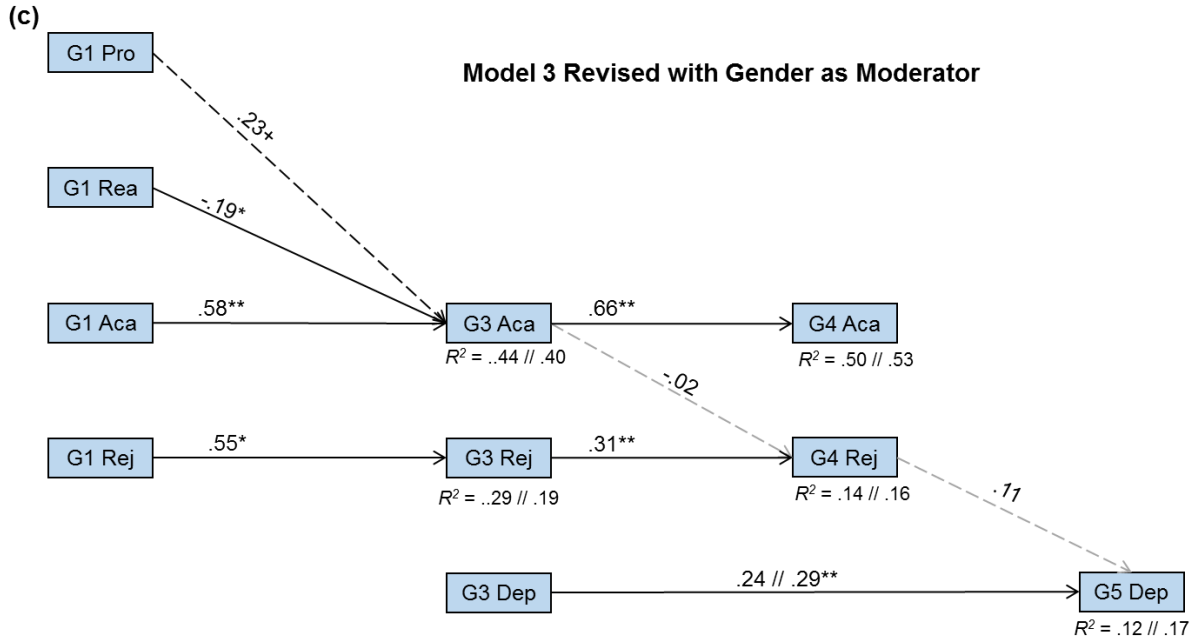


Figure 5. Results of four alternative path models revised with gender as moderator

(Continues on next page.)



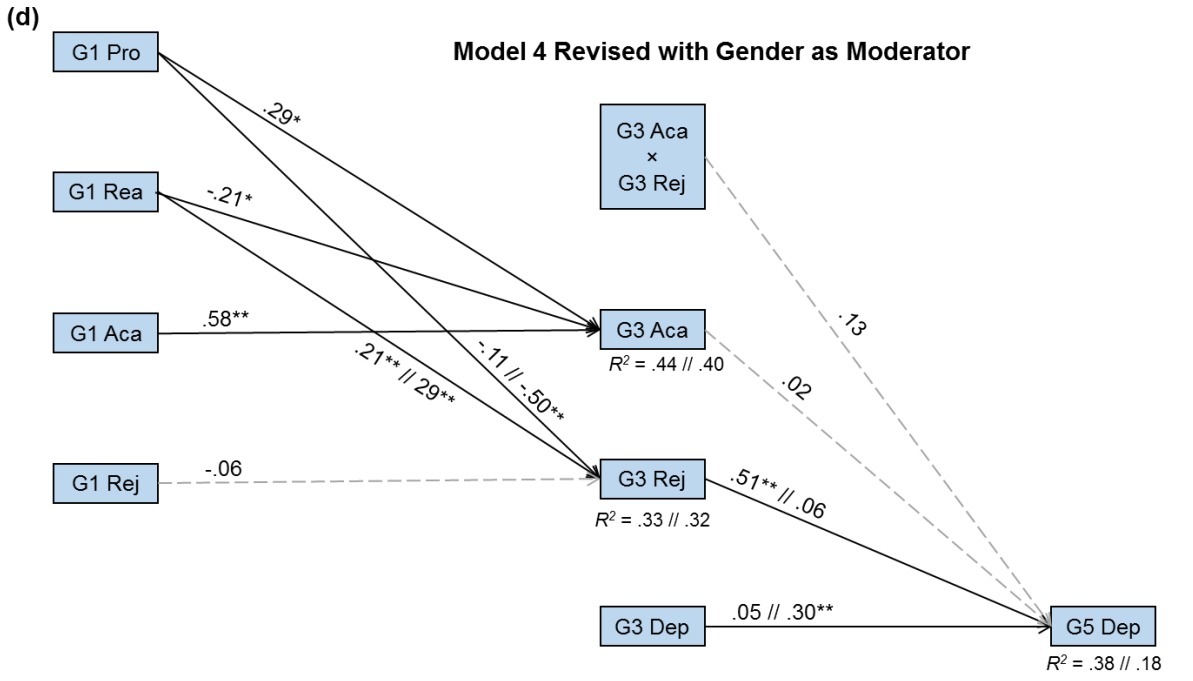


Figure 5. Results of four alternative path models revised with gender as moderator
Gender-conditional effects are reported as boys // girls. Unconditional effects are reported as a single estimate for the whole sample. Black lines denote paths that are significant at $p < .05$; gray dashed lines denote nonsignificant paths ($p \geq .05$). Estimates are unstandardized. + $p < .1$, * $p < .05$, ** $p < .01$.

Model Evaluation and Refinement

Model fit statistics for all models reported above are presented in Table 4. Comparing Model 2 and 3 (non-nested but directly comparable), AIC, BIC, and sample-adjusted BIC values all show slight improvements of Model 2 (social cascade model) over Model 3 (academic cascade model). This pattern of results remains the same in the single-group and gender-moderated group models (Model 2b vs. 3b). Further, given the meaningful significant moderation of the path from third grade peer rejection to fourth grade academic performance in Model 2 and the absence a significant path from third grade academic performance to fourth grade peer rejection in Model 3, Model 2 also appears to be more substantively and theoretically meaningful.

With respect to comparing model 1 and 4, the absence of an interaction between academic performance and peer rejection in Model 4 renders the entire model unnecessary. Thus, Model 1 is preferred over Model 4.

Given these results, Models 1 and 2 were refined to create a single model to account for all of the significant developmental pathways found in the present study. Results of the final model are presented in Figure 6. Fit statistics are presented in Table 4, first for the model with all paths retained from Models 1 and 2 (Model 5a), and then for the model with nonsignificant cross-lagged paths trimmed for parsimony (Model 5b, Figure 6). Thus, Model 5b is considered the final model.

As shown in Table 3, the final model exhibited improved fit in terms of AIC, BIC, and SABIC fit statistics, relative to the previous model. While RMSEA and chi square ($\chi^2/df = 2.81$) statistics suggested adequate fit, other indicators (CFI = .843, TLI = .760, SRMR = .106) indicated that overall model fit could be improved, particularly by estimating covariance paths

among concurrently measured variables. However, this was not pursued considering it was not a primary component of the analytic plan and strategies for evaluating models.

The final model demonstrates that proactive aggression in first grade predicts better academic performance in third grade for boys and girls as well as lower levels of peer rejection in third grade for girls only. Conversely, reactive aggression in first grade predicts poorer academic performance and higher levels of peer rejection for both boys and girls, with the effects of reactive aggression on peer rejection being particularly pronounced for girls. Gender differences become more pervasive in the later developmental pathways spanning third through fifth grade. For boys but not girls, peer rejection in third grade predicts poor academic performance in fourth grade as well as depressive symptoms in fifth grade. For girls but not boys, peer rejection in third grade may actually predict better academic performance in fourth grade (although this path was only marginally statistically significant), with no association with fifth grade depressive symptoms; rather, girls' depressive symptoms were accounted for by earlier depressive symptoms only (this stability path was not found to be significant for boys). Finally, academic performance, but not peer rejection, was stable from first to third grade, and both were stable from third to fourth grade.

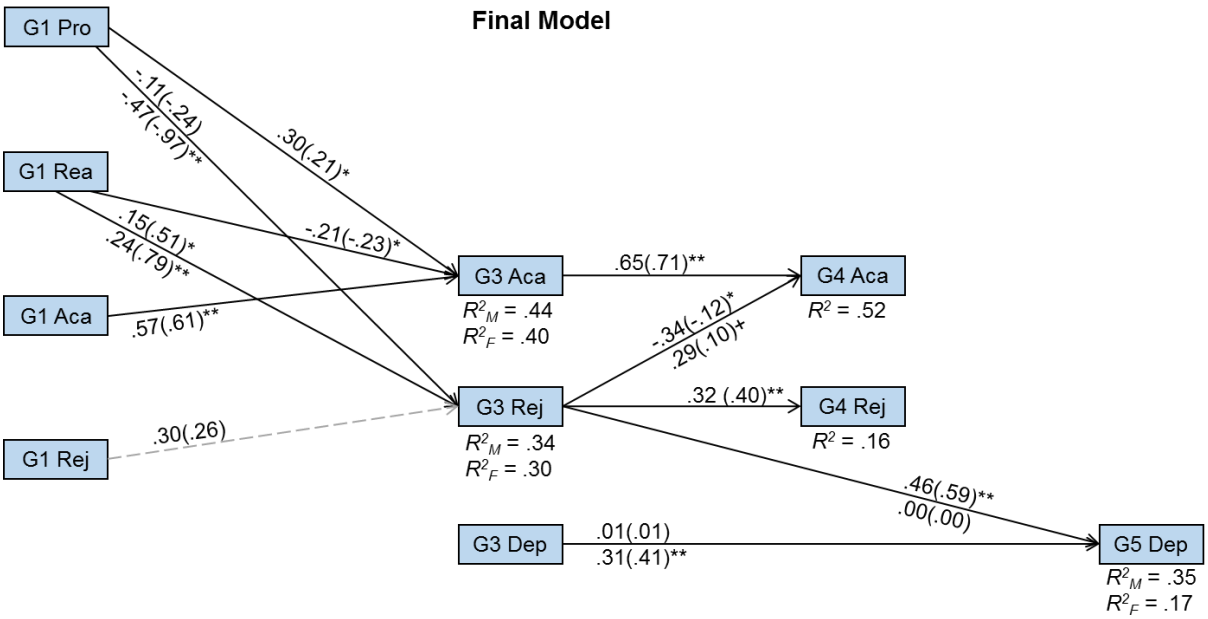


Figure 6. Results of the final revised model
Where effects are moderated by gender, the estimate for boys appears above the line and the estimate for girls below. Black lines denote paths that are significant at $p < .05$; gray dashed lines denote nonsignificant paths ($p \geq .05$). Unstandardized path coefficients are reported outside parentheses, with standardized estimates inside parentheses. + $p < .1$, * $p < .05$, ** $p < .01$.

Table 6. Parameter estimates and standard errors from final model

	Gender Differences						No Gender Differences		
	Boys			Girls			Full Sample		
	B	S.E.	β	B	S.E.	β	B	S.E.	β
G5 Dep on									
G3 Dep	0.005	0.129	0.006	0.306**	0.109	0.409	-	-	-
G3 Rej	0.460**	0.132	0.586	0.002	0.088	0.002	-	-	-
G4 Aca on									
G3 Aca	-	-	-	-	-	-	0.649**	0.040	0.694
G3 Rej	-0.335*	0.163	-0.121	0.291+	0.151	0.098	-	-	-
G4 Rej on									
G3 Rej	-	-	-	-	-	-	0.323**	0.091	0.385
G3 Aca on									
G1 Aca	-	-	-	-	-	-	0.573**	0.061	0.610
G1 Pro	-	-	-	-	-	-	0.300*	0.136	0.233
G1 Rea	-	-	-	-	-	-	-0.212*	0.087	-0.251
G3 Rej on									
G1 Rej	-	-	-	-	-	-	0.300	0.223	0.287
G1 Pro	-0.106	0.118	-0.244	-0.470**	0.148	-0.971	-	-	-
G1 Rea	0.146*	0.059	0.510	0.240**	0.080	0.792	-	-	-
R-Squares									
G5 Dep	0.345*	0.165	-	0.168+	0.101	-	-	-	-
G4 Aca	0.523**	0.056	-	0.523**	0.057	-	-	-	-
G4 Rej	0.148+	0.087	-	0.163*	0.072	-	-	-	-
G3 Aca	0.442**	0.060	-	0.395**	0.067	-	-	-	-
G3 Rej	0.341**	0.107	-	0.297*	0.134	-	-	-	-

+ $p < .1$, * $p < .05$, ** $p < .01$.

Discussion

The present study examined four variations of the failure model to determine the roles of peer rejection and academic performance in the developmental pathways from aggression to depressive symptoms. As originally hypothesized (Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991), the model specified that peer rejection and academic performance might constitute dual pathways through which early aggression leads to later depression; however, the relative magnitude, temporal sequence, and possible interaction of these two variables had not been clearly articulated or investigated. Further, the differential effects of reactive and proactive aggression had not been considered. The results of the present study provide partial support for the failure model as a whole and lend some clarity to the developmental sequences and through which this developmental process might occur. The primary findings are as follows.

With respect to the early pathways from aggression in first grade to academic and social functioning in third grade, (a) reactive aggression predicted subsequent lower levels of academic performance and higher levels of peer rejection for both boys and girls, with the link between reactive aggression and peer rejection being strongest for girls; however, (b) proactive aggression predicted subsequent higher levels of academic performance for both boys and girls, and lower levels of peer rejection for girls only. With respect the later pathways from academic and social functioning in third grade to depressive symptoms in fifth grade, (c) for boys only, peer rejection predicted subsequent poor academic performance and depressive symptoms; and (d) for girls only, earlier depressive symptoms predicted subsequent depressive symptoms, but peer rejection and academic performance did not. Further, (e) third grade peer rejection and academic performance did not interact (i.e., no buffering or exacerbating effect) in predicting fifth grade depressive symptoms. Overall, these results provide some preliminary support for the

model's peer rejection pathway and social cascade hypothesis, particularly among boys, but no evidence was found for academic performance as a predictor of, or mediating pathway to, depressive symptoms. These results are discussed and interpreted in further detail below.

Social and Academic Sequelae of Early Proactive and Reactive Aggression

Previous research has established associations between reactive aggression and both social (e.g., Card & Little, 2006; Salmivalli & Helteenvuori, 2007) and academic (e.g., Day, Beam, & Pal, 1992; Fite et al., 2013) problems. The present results accord with and expand upon these prior findings. Notably, the standardized coefficients between reactive aggression and peer rejection were more robust and of at least twice the magnitude of those between reactive aggression and academic performance, suggesting that reactive aggression may be more detrimental for social functioning than for academic functioning. During the preschool through early elementary school years, the most observable forms of aggressive behavior (i.e., physical aggression, whether proactive or reactive in nature) typically decline for most children (Vitaro & Brendgen, 2012). Children who continue to exhibit reactively aggressive behaviors more than their peers during first grade and beyond may be at risk for significant social problems, which in the present study manifested two years later in the form of third grade peer rejection. It may be that reactively aggressive behaviors interfere with friendship formation and social skill development, leading children to experience compounding social problems over time.

Interestingly, this link between reactive aggression and peer rejection was moderated by gender, with reactive aggression being more strongly associated with peer rejection for girls than for boys. This finding makes some sense, as evidence suggests that aggression is more strongly linked to psychosocial problems when the types of aggressive behaviors are inconsistent with gender norms (Crick, 1997). That is, because aggressive behaviors are more normative in boys

than in girls (as indicated by previous research and consistent with the present results), it may be that aggressive behaviors committed by girls, as compared to boys, are considered more abnormal and less acceptable by same-age peers, leading to greater peer rejection (Crick, 1997).

While reactive aggression confers greater risk for peer rejection for girls, proactive aggression appears to follow the opposite pattern, predicting significantly lower levels of peer rejection for girls but not for boys. Previous research has been inconsistent with regard to the association between proactive aggression and social problems in youth, particularly compared to and after controlling for the association between reactive aggression and social problems (Card & Little, 2006). For example, Fite and colleagues (Evans et al., 2015; Fite et al., 2012b, 2013), studying different samples of children and adolescents, have found that significant zero-order associations between proactive aggression and peer rejection are attenuated to nonsignificance after controlling for the effects of reactive aggression. However, these studies have not provided clear longitudinal tests of gender differences in the association between proactive aggression and social problems. The present findings raise the possibility that there may be gender-specific associations between proactive aggression and peer rejection, but these may have been obfuscated by mixed samples and male-only samples in previous research.

The negative association found between reactive aggression and academic performance is consistent with previous research (e.g., Day et al., 1992; Fite et al., 2013). However, the positive association between proactive aggression and academic performance was unexpected and inconsistent with previous research demonstrating no unique association (e.g., Fite et al., 2013) or a negative association (e.g., Day et al., 1992). While these results should be interpreted with caution, they also underscore the utility of delineating the proactive and reactive functions of aggression. Reactively aggressive children are more impulsive and likely to have a range of

cognitive, social, and behavioral difficulties, whereas proactively aggressive youth are often somewhat successful in various aspects of social-emotional functioning (Fite et al., 2012a; Vitaro & Brendgen, 2012). Given the planful, premeditative nature of proactive aggression, proactively aggressive youth may have greater cognitive and self-regulatory resources (Fite et al., 2012a; Vitaro & Brendgen, 2012). These abilities may equip some children both to commit proactively aggressive acts and to succeed academically. In other words, the characteristics and correlates associated with reactive aggression might be academically detrimental whereas the opposite may be true for proactive aggression.

Finally, it should be acknowledged that some of the gender differences in pathways from early aggression subtypes to subsequent academic and social problems might be better understood by distinguishing between the forms of aggression (physical vs. relational), in addition to the functions of aggression that were considered here. Research has documented that relational aggression (e.g., deliberately excluding a peer from a social activity) can be observed in children as early as the preschool years (Ostrov & Crick, 2005), and it can continue or increase throughout childhood and adolescence even as physical aggression declines (Vitaro & Brendgen, 2012). Relational aggression appears to play a particularly important role for girls, manifesting both proactively and reactively, with mixed effects on social functioning (Crapanzano, Frick, & Terranova, 2010; Rys & Bear, 1997; Smith, Rose, & Schwartz-Mette, 2009). For example, findings from different studies conducted with early childhood samples suggest that relationally aggressive girls may have more friends, lower peer acceptance, higher expressive language skills, and more often receive a “controversial” status by their peers (i.e., receiving both high and low sociometric ratings; Ostrov & Crick, 2005). Thus, considering that at least some of these studies have found early relational aggression to be associated with

positive social correlates and outcomes among girls, this may help explain the present finding that, for girls only, early proactive aggression predicted lower subsequent levels of peer rejection. Clearly more research is needed to better understand the developmental sequelae of form-and-function subtypes of aggression among boys and girls.

Social and Academic Pathways to Depressive Symptoms

The failure model hypothesizes academic and social "failure" as two mediating pathways leading to depressive symptoms. Results provide some support was for the social pathway, for boys only. This finding is consistent with the original articulation and interpretations of the failure model, although the evidence was limited. First, Patterson's failure model was originally tested among samples of boys (Patterson & Capaldi, 1990; Patterson & Stoolmiller, 1991), but there was nothing in the theoretical model that specified why it would only apply to boys and not to girls. Second, both of these early studies concluded that the social pathway was stronger or more consistent than the academic pathway, but this was based on cross-sectional results drawn from high-risk samples of boys.

The present findings suggest that boys' peer relationships play an important role during middle childhood, perhaps more so than academic performance and perhaps more so than peer relationships during other periods. Developmentally, third grade is a time during which children's peer networks are established and expanding, with peer influences playing an increasingly larger role in their lives into adolescence (Eccles, 1999). In contrast, academically, third grade marks a transition from acquisition to application of basic skills such as reading, which serves as a robust predictor of academic outcomes throughout primary and secondary education (Hernandez, 2011). Children who are falling behind academically may have much greater difficulty during the late elementary and secondary school years, as they fall farther and

farther behind relative to their peers. Thus, to the extent that academic "failure" could be a risk factor for depressive symptoms, it may be that children have not yet accumulated the quality and quantity of academic problems to experience that risk. Although this study and previous studies have primarily considered the relative risk of academic vs. social problems during the same limited time frames, these speculations raise an intriguing alternative possibility: perhaps both pathways contribute to depressive symptoms, but during different developmental periods (e.g., peer rejection during middle childhood, academic performance during adolescence). Further research will be needed to explore this possibility and to replicate the present findings.

The gender differences in the paths to depressive symptoms are not surprising in light of past research conducted among older samples. For example, McCarty et al. (2008) found that academic failure predicted subsequent depressive episodes for adolescent girls but not boys, whereas Obradovic et al. (2010) found that social competence predicted depressive symptoms during emerging adulthood for males but not for females. Thus, although the peer rejection to depressive symptoms pathway was significant for boys and not for girls, this should not be interpreted as evidence that girls are at lower risk for depressive symptoms or that peer rejection is less consequential for girls during these periods. Indeed, in the present sample across all time points, gender was completely uncorrelated with depressive symptoms and showed only weak inconsistent associations with peer rejection (Table 2). This supports the notion that the present findings differ from what might be found among an older sample, as gender differences in depressive symptoms are negligible during childhood and then become more pronounced as youth enter and develop through adolescence (Hankin et al., 1998; Twenge & Nolen-Hoeksema, 2002). While many explanations have been proposed that may help gender differences in depressive symptoms, evidence seems to support some form of transactional or diathesis-stress

models of psychopathology. Compared to their male counterparts, early-adolescent girls appear to exhibit higher average levels of certain psychological and interpersonal characteristics (e.g., dependence on others, stress reactivity, agreeableness), which may be helpful in some contexts but also increase risk for internalizing problems. These characteristics, in turn, may increase vulnerability to the depressogenic effects of stressful life events (particularly interpersonal stressors), which are prevalent among adolescent girls (Leadbeater, Blatt, & Quinlan, 1995; Nolen-Hoeksema & Girgus, 1994; Oldehinkel & Bouma, 2011). However, the present study did not assess girls during the developmental period in which such diathesis-stress patterns would be expected. Instead, a possible explanation for the observed gender differences may be that girls simply showed higher stability, compared to boys, in their level of depressed symptoms or withdrawn behavior from third to fifth grade, and this might be related to gender-specific social, personality, or rater factors. This stability of depressive symptoms observed in girls may have obfuscated any associations between third grade predictors and fifth grade depressive symptoms, whereas the low stability of depressive symptoms observed in boys may have allowed for more variance in fifth grade depressive symptoms to be accounted for by earlier peer rejection. This possibility further underscores the need for cautious interpretation and future replication.

Finally, even though the cascade sequence models (Models 2 and 3) were not supported in terms of leading to depressive symptoms, results did yield the unexpected finding that peer rejection in third grade predicted poor academic performance in fourth grade among boys but not among girls. Further, there was a trend in the opposite direction for girls, suggesting that peer rejection might predict *better* academic performance. Thus, the present results, if replicated, may indicate that peer rejection during middle childhood might be particularly detrimental boys, affecting them in social, academic, and affective domains. These findings make some sense in

light of previous research documenting gender-specific protective factors, such as friendship quality, which serves to buffer the risk between victimization and depression and anxiety for girls but not for boys (Schmidt & Bagwell, 2007). In the absence of such protective factors, boys who experience peer rejection may be at higher risk for poor outcomes, such as internalizing problems. For boys, being rejected by peers may interfere with their ability to perform well in school. Alternatively, there may be a third factor, such as hyperactive-impulsive behaviors, which occur more frequently in boys and could account for associations among reactive aggression, poor academic performance, and peer rejection (Evans et al., 2015; McConaughy et al., 2011).

It is not clear why girls would show a marginally statistically significant positive path from third grade peer rejection to fourth grade academic performance, whereas boys demonstrated the opposite (a negative path). It may be the case that girls who have difficulty with some peers are able to continue at their usual level of academic performance or better, whereas boys may experience more pervasive problems as a result. Alternatively, these differences may be linked to gender differences in the social structures of boys and girls friendships in middle childhood. Specifically, it is thought that girls tend to have closer dyadic friendships during middle childhood whereas boys tend to have friendships based on shared activities and interests at a broader group level (Underwood, 2007). Accordingly, for boys, peer rejection may represent both a broader pattern of social exclusion as well as an absence of supportive friendships to help buffer those effects.

Limitations and Future Directions

Several limitations of the present study should be noted. First, assessments were entirely by teacher report. Although previous research supports the validity of teachers' ratings of

proactive and reactive aggression (e.g., Dodge & Coie, 1987), peer rejection and academic performance (e.g., Evans et al., 2015, 2016; Fite et al., 2013), and depressive symptoms (e.g., Achenbach & Rescorla, 2001), the present findings are nonetheless susceptible to mono-informant bias. Future research should incorporate data from multiple sources. Self-report and other assessments of depressive symptoms would be particularly helpful toward measuring the full range of depressive symptoms in youth (Klein, Dougherty, & Olino, 2005). Further, multi-informant/-method assessment might also facilitate better detection and explanation of depressive symptoms in girls, which appears to be a particular limitation of the study data. Related to the issue of teacher-report bias is that of students being nested within different classrooms at different time points. While the effects of teacher- and cohort-related dependencies seem minimal (see Appendix B), they do raise the possibility of Type II errors given the unaccounted-for, presumably random variance they introduce into the data.

Further, the study design and analytic plan did not permit comprehensive examination of all of the types of questions that might be of interest in longitudinal mediation models. Specifically, developmental pathways were examined from one time point to another without examination of the patterns of intra-individual change over time. Future research would benefit from closer examination of trajectories of change (via latent growth curve, multilevel, or growth mixture modeling) in proactive and reactive aggression, peer rejection and academic performance, and depressive symptoms. With respect to the question of mediation, the accelerated design only allowed estimation of indirect paths only (e.g., reactive aggression to peer rejection to depressive symptoms) and direct paths (e.g., reactive aggression to depressive symptoms) could not be considered. Research is needed to examine true mediational models in which at least one cohort of individuals (preferably more) is followed from the T1 predictor, T2

mediator, and T3 outcome variable. Another consequence of the accelerated design is the high rate of planned missing data, which may have inflated standard errors and increased the likelihood of a Type II error. For this reason, the present findings should be interpreted as preliminary and in need of replication in future research.

An additional limitation lies in the restricted range of variables that were assessed in the present study. For example, the proactive and reactive functions of aggressive behavior were assessed, but the *forms* of aggressive behavior (e.g., physical vs. relational aggression) were not examined. Similarly, researchers have conceptualized the social, academic, and emotional components of the failure model in a variety of ways (e.g., social competence rather than peer rejection, internalizing problems rather than depressive symptoms). While the constructs measured in this study were selected and measured specifically for their hypothesized role in the model based on previous research, there are likely other facets or subtypes of these variables that should also be considered. Delineating function and form of aggression may be particularly important given that the Dodge and Coie (1987) measure used in this study could be interpreted by some as pulling more strongly for physical-proactive and physical-reactive aggression than as being “neutral” on the form dimension.

Finally, it should be reiterated that the present study was conducted among a predominately Caucasian sample of school-age children, from a single school. This limits the generalizability of the findings in a number of ways. Future research is needed to examine these questions across youth of diverse ethnic and socioeconomic backgrounds, as well as children who are at risk in various domains. In particular, it might be useful to examine outcomes involving depressive symptoms and antisocial behavior during adolescence when rates of such

behavioral problems increase. Similarly, future research should examine these developmental models among at-risk populations, such as clinical and adjudicated samples.

Conclusions and Implications

The present study provided a direct longitudinal examination of four variations of the failure model, while also incorporating the distinct roles of proactive and reactive aggression. The results provided partial support for the failure model as an explanation for the developmental pathways from aggression to depressive symptoms. Specifically, reactively aggressive first graders were at risk for subsequent peer rejection and poor academic performance in third grade. From that point, boys who were rejected by peers were more likely to go on to experience depressive symptoms in fifth grade. These findings shed light on the failure model in three respects. First, results suggest that the social-failure pathway seems more tenable than the academic-failure pathway (and not simply because the academic pathway has received less empirical attention). Second, in the present sample it was largely boys, not girls, who progressed from peer rejection to depressive symptoms, suggesting that this model may fit better for boys than for girls. Finally, distinguishing between the proactive and reactive functions of aggression does appear to improve the model. In general, it seems to be reactive aggression that confers risk for poorer academic and social outcomes among both boys and girls, adding to a body of evidence pointing to reactive aggression as a risk factor for multiple poor psychosocial outcomes (Fite et al., 2012a). There is a need for further research to replicate the present findings and to further examine the social pathways from reactive aggression to depressive symptoms as well as alternative developmental mechanisms for girls and for the effects of poor academic performance.

This investigation also highlights specific directions for future research, the most central of which is the need for independent replication in a longitudinal study with greater data coverage. As noted in the limitations section, research is needed in order to establish true longitudinal mediation; to improve measurement via multiple informants and methods; to consider alternative developmental time period and timing of measurements; and to investigate alternative populations, such as clinically referred youth and samples with greater diversity in cultural, ethnic, and socio-economic backgrounds. While these are all important directions for future research in a general sense, there are also some specific research questions that warrant direct hypothesis testing.

First, the present study focused specifically on the *functions* of aggressive behavior (i.e., proactive and reactive aggression), but aggression also takes different *forms* which are theoretically, empirically, and practically meaningful in terms of correlates, etiologies, and outcomes (Little et al., 2003; Vitaro & Brendgen, 2012). Relational aggression (e.g., deliberately excluding a peer from group activities) and physical aggression (e.g., hitting or pushing a peer) take different forms, occur in different settings, may not be monitored and managed in the same way. In theory, forms and functions of aggression are orthogonal to one another, such that both physically and relationally aggressive acts could be enacted to serve proactively or reactively aggressive functions. The child aggression literature provided specific reason for this investigation to focus on functions of aggressive behavior, but form of aggression may also play a role. Research examining the failure model from a form-by-function approach could help clarify even more precisely what kind of aggressive behavior confers risk for developmental pathways leading to social, academic, and mood problems. This might be particularly helpful for clarifying the gender differences observed in the present study, given that relational aggression

appears to play a particularly important role among girls while physical and reactive aggression are more common in boys (Vitaro & Brendgen, 2012).

Second and more broadly, there is a need for research to test the failure model from alternative theoretical and analytic approaches. The present design utilized a minimalist, fine-grained model comparison and model-building approach, which could also be considered a variable-centered design. Other variable-centered designs, such as fully cross-lagged models, could help elucidate the observed paths among variables from a more empirical (as opposed to theory-driven) perspective. There is also a great deal to learn from person-centered designs, such as growth mixture and latent growth curve models which can differentiate individuals or subgroups of children based on their trajectories of aggression, peer rejection, academic performance, and/or depressive symptoms. While such studies might yield findings similar to or different from those of the present study, they essentially ask different questions and therefore can be a useful complement to variable-oriented analyses.

In addition, these findings have practical implications for assessment and intervention. Given its centrality as a social-emotional and academic risk factor, reactive aggression appears to be an important target for early screening and intervention. This study examined aggression in first grade because typically physical aggression declines after the socialization that occurs in preschool and kindergarten. Children who continue to be more reactively aggressive than their peers in first and second grade should raise flags, not only to address the aggression, but also to closely monitor their social-emotional and academic trajectories. School-based interventions targeting anger and aggressive behavior, such as Coping Power (Lochman & Wells, 2004), may be beneficial in curbing early reactive aggression. However, many of these interventions have generally been developed and tested among older elementary school and middle school age

youth. The present findings demonstrate the need for earlier intervention, which is less likely to involve child-focused skill-building approaches and more likely to involve behavioral interventions targeted at parents and teachers, such as Incredible Years, Parent Child Interaction Therapy, Triple-P, and Teacher Child Interaction Training (Fernandez, Gold, Hirsch, & Miller, 2015; Thomas & Zimmer-Gembeck, 2007; Webster-Stratton, Reid, & Stoolmiller, 2008). By reducing disruptive behavior and improving social-emotional functioning during the early elementary school years, mental health and school professionals can help reduce aggression and build a foundation for academic and social success.

With regard to risk for depressive symptoms, these findings suggest that during later elementary school, boys and girls may follow different paths. For boys, peer rejection in third grade appears to be a particularly pivotal issue, which can predict higher levels of depressive symptoms two years later. Thus, boys who are excluded, teased, disliked, and victimized, particularly in the absence of any close friends, may benefit from multifaceted interventions that bolster social skills, self-esteem, and opportunity for building friendships in well-supervised environments (e.g., DeRosier, 2007), while also involving materials specifically designed for preventing depressive and internalizing problems (e.g., Barrett, Farrell, Ollendick, & Dadds, 2006). It should be reiterated that for the vast majority of children, depressive symptoms were within the subclinical range, but these might be considered as prodromal or subthreshold risk for subsequent depression during adolescence.

Finally, it would be inappropriate to conclude from these results that girls are somehow at lower risk for depressive symptoms. Rather, the stability of depressive symptoms from third to fifth grade among girls but not boys is remarkable and suggests that girls may experience prolonged depressive or withdrawn symptoms more persistently than boys. The finding that

girls' depressive symptoms were not predicted by academic or social problems suggests that this study did not examine or detect antecedents to girls' depressive symptoms. Despite the possibility of different mechanisms, both girls and boys may benefit from early preventive interventions for depressive symptoms (e.g., Barrett et al., 2006).

These findings help advance the theoretical understanding of developmental pathways and our knowledge of *when* and *for whom* interventions are likely to be most helpful. Future research and clinical efforts are needed to further validate these and other developmental pathways, and to translate such findings into effective assessment and intervention efforts.

References

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

Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Barker, E. D., & Salekin, R. T. (2012). Irritable oppositional defiance and callous unemotional traits: Is the association partially explained by peer victimization? *Journal of Child Psychology and Psychiatry, 53*, 1167–1175.

Barrett, P. M., Farrell, L. J., Ollendick, T. H., & Dadds, M. (2006). Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: An evaluation of the FRIENDS Program. *Journal of Clinical Child and Adolescent Psychology, 35*, 403-411.

Barthelemy, J. J., & Lounsbury, J. W. (2009). The relationship between aggression and the big five personality factors in predicting academic success. *Journal of Human Behavior in the Social Environment, 19*, 159-170.

Becker, S. P., Fite, P. J., Vitulano, M. L., Rubens, S. L., Evans, S. C., & Cooley, J. L. (2014). Examining anxiety and depression as moderators of the associations between ADHD symptoms and academic and social problems in Hispanic adolescents. *Journal of Psychopathology and Behavioral Assessment*, *36*, 265-275.

Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin, 106*, 59–73.

Boots, D. P., Wareham, J., & Weir, H. (2011). Gendered perspectives on depression and antisocial behaviors: An extension of the failure model in adolescents. *Criminal Justice and Behavior*, *38*, 63-84.

Brennan, L. M., Shaw, D. S., Dishion, T. J., & Wilson, M. (2012). Longitudinal predictors of school-age academic achievement: Unique contributions of toddler-age aggression, oppositionality, inattention, and hyperactivity. *Journal of Abnormal Child Psychology, 40*, 1289-1300.

Burke, J. D., Loeber, R., Lahey, B. B., & Rathouz, P. J. (2005). Developmental transitions among affective and behavioral disorders in adolescent boys. *Journal of Child Psychology and Psychiatry, 46*, 1200-1210.

Campbell, S. B., Spieker, S., Burchinal, M., & Poe, M. D. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry, 47*, 791-800.

Capaldi, D. M. (1992). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: II. A 2-year follow-up at grade 8. *Development and Psychopathology, 4*, 125-144.

Card, N., & Little, T. (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.

Chen, X., Huang, X., Chang, L., Wang, L., & Li, D. (2010). Aggression, social competence, and academic achievement in Chinese children: A 5-year longitudinal study. *Development and Psychopathology, 22*, 583-592.

Cicchetti, D., & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology, 70,* 6-20.

Coie, J., Terry, R., Lenox, K., & Lochman, J. (1995). Childhood peer rejection and aggression as predictors of stable patterns of adolescent disorder. *Development and Psychopathology, 7,* 697-713.

Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology, 112,* 558-577.

Crapanzano, A. M., Frick, P. J., & Terranova, A. M. (2010). Patterns of physical and relational aggression in a school-based sample of boys and girls. *Journal of Abnormal Child Psychology, 38,* 433-445.

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

Crick, N. R., Ostrov, J. M., & Werner, N. E. (2006). A longitudinal study of relational aggression, physical aggression, and children's social-psychological adjustment. *Journal of Abnormal Child Psychology, 34,* 131-142.

Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1999). Problems reported by parents and children in multiple cultures: The Cchild Bbehavior Cchecklist syndrome constructs. *The American Journal of Psychiatry, 156,* 569-74.

Day, D. M., Bream, L. A., & Pal, A. (1992). Proactive and reactive aggression: An analysis of subtypes based on teacher perceptions. *Journal of Clinical Child Psychology, 21,* 210-217.

DeRosier, M. E. (2007). Peer-rejected and bullied children: A Safe Schools Initiative for elementary school students. In J. E. Zins, M. J. Elias, & C. A. Maher (Eds.), *Bullying, victimization, and peer harassment: A handbook of prevention and intervention* (pp. 257-271). Binghamton, NY: Haworth Press.

Dishion, T. J., Véronneau, M., & Myers, M. W. (2010). Cascading peer dynamics underlying the progression from problem behavior to violence in early to late adolescence. *Development and Psychopathology, 22*, 603-619.

Dodge, K. A., & Coie, J. D. (1987). Social information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53*, 1146-1158.

Dodge, K. A., Lochman, J. E., Harnish, J. D., Bates, J. E., & Pettit, G. S. (1997). Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. *Journal of Abnormal Psychology*, *106*, 37-51.

Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological Methods, 1*, 236-248.

Eccles, J. S. (1999). The development of children ages 6 to 14. *The Future of Children*, *9*, 30-44.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*, 430-457.

Evans, S. C., Fite, P. J., Hendrickson, M. L., Rubens, S. L., & Mages, A. K. (2015). The role of reactive aggression in the link between hyperactive-impulsive behaviors and peer rejection in adolescents. *Child Psychiatry and Human Development*, *46*, 903-912.

Evans, S. C., Pederson, C. A., Fite, P. J., Blossom, J. B., & Cooley, J. L. (2016). Teacher-reported irritable and defiant dimensions of Oppositional Defiant Disorder: Social, behavioral, and academic correlates. *School Mental Health*, *8*, 292-304.

Fernandez, M. A., Gold, D. C., Hirsch, E., & Miller, S. P. (2015). From the clinics to the classrooms: A review of teacher-child interaction training in primary, secondary, and tertiary prevention settings. *Cognitive and Behavioral Practice*, *22*, 217-229.

Firth, S. M., Shohet, K. L., & Thurber, S. (2000). An extension of the dual failure model of depression to adults. *North American Journal of Psychology*, *2*, 403-408.

Fite, P. J., Colder, C. R., & Pelham, W. E. (2010). A factor-analytic approach to distinguish pure and co-occurring dimensions of proactive and reactive aggression. *Journal of Clinical Child and Adolescent Psychology, 35*, 578- 582.

Fite, P. J., Hendrickson, M., Rubens, S., Gabrielli, J., & Evans, S. (2013). The role of peer rejection in the link between reactive aggression and academic performance. *Child and Youth Care Forum*, *42*, 193-205.

Fite, P. J., Rathert, J., Colder, C. R., Lochman, J. E., & Wells, K. C. (2012a). Proactive and reactive aggression. In R. J. R. Levesque (Ed.), *Encyclopedia of adolescence* (pp. 2164–2170). New York, NY: Springer.

Fite, P. J., Rubens, S. L., Preddy, T. M., Raine, A., & Pardini, D. A. (2014). Reactive/proactive aggression and the development of internalizing problems in males: The moderating effect of parent and peer relationships. *Aggressive Behavior*, *40*, 69-78.

Fite, P. J., Stoppelbein, L., & Greening, L. (2009). Proactive and reactive aggression in a child psychiatric inpatient population. *Journal of Clinical Child and Adolescent Psychology*, *38*, 199-205.

Fite, P. J., Wimsatt, A. R., Vitulano, M. L., Rathert, J. L., & Schwartz, S. (2012b). Examination of peer rejection and depressive symptoms as mediators of the link between rule-breaking behavior and poor academic performance. *Journal of Psychopathology and Behavioral Assessment, 34,* 164-171.

Frick, P. J., & Viding, E. (2009). Antisocial behavior from a developmental psychopathology perspective. *Development and Psychopathology, 21*, 1111-1131.

Gooren, E. M., van Lier, P. A., Stegge, H., Terwogt, M. M., & Koot, H. M. (2011). The development of conduct problems and depressive symptoms in early elementary school children: The role of peer rejection. *Journal of Clinical Child & and Adolescent Psychology, 40*, 245-253.

Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: a A dynamic systems approach. *Psychological Review*, *113*, 101-131.

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology*, *107*, 128-140.

Hernandez, D. J. (2011). Double jeopardy: How third-grade reading skills and poverty influence high school graduation. *Annie E. Casey Foundation*.

Hox, J. J. (2002). *Mutlilevel analysis: Techiques and applications.* Mahwah, NJ: Erlbaum.

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.

Huesmann, L. R., Dubow, E. F., & Boxer, P. (2009). Continuity of aggression from childhood to early adulthood as a predictor of life outcomes: Implications for the adolescent-limited and life-course-persistent models. *Aggressive Behavior, 35*, 136-149.

Ingoldsby, E. M., Kohl, G. O., McMahon, R. J., Lengua, L., & The Conduct Problems Prevention Research Group. (2006). Conduct problems, depressive symptomatology, and co-occurring presentation in childhood as predictors of adjustment in early adolescence. *Journal of Abnormal Child Psychology, 34*, 603-621.

Kelly, J. B., & Emery, R. E. (2003). Children's adjustment following divorce: Risk and resilience perspectives. *Family Relations, 52*, 352-362

Klein, D. N., Dougherty, L. R., & Olino, T. M. (2005). Toward guidelines for evidence-based assessment of depression in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*, 412-432.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford.

Leadbeater, B. J., Blatt, S. J., & Quinlan, D. M. (1995). Gender-linked vulnerabilities to depressive symptoms, stress, and problem behaviors in adolescents. *Journal of Research on Adolescence*, *5*, 1-29.

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

Little, T. D., Henrich, C. C., Jones, S. M., & Hawley, P. H. (2003). Disentangling the “whys” from the “whats” of aggressive behaviour. *International Journal of Behavioral Development, 27*, 122-133.

Lochman, J. E., & Wells, K. C. (2004). The Coping Power Program for preadolescent boys and their parents: Outcome effects at the 1-year follow-up. *Journal of Consulting and Clinical Psychology*, *72*, 571-578.

Loth, A. K., Drabick, D. A. G., Leibenluft, E., & Hulvershorn, L. A. (2014). Do childhood externalizing disorders predict adult depression? A meta-analysis. *Journal of Abnormal Child Psychology*. *42*, 1115–1116.

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.

Marmorstein, N., R., & Iacono, W. G. (2011). Explaining associations between cannabis use in adolescence and later major depression: A test of the psychosocial failure model. *Addictive Behaviors*, *36*, 773-776.

Mathieson, L. C., & Crick, N. R. (2010). Reactive and proactive subtypes of relational and physical aggression in middle childhood: Links to concurrent and longitudinal adjustment. *School Psychology Review*, *39*, 601-611.

McCarty, C. A., Mason, W. A., Kosterman, R., Hawkins, J. D., Lengua, L. J., & McCauley, E. (2008). Adolescent school failure predicts later depression among girls. *Journal of Adolescent Health*, *43*, 180-187.

McConaughy, S. H., Volpe, R. J., Antshel, K. M., Gordon, M., & Eiraldi, R. B. (2011). Academic and social impairments of elementary school children with attention deficit hyperactivity disorder. *School Psychology Review*, *40*, 200-225.

Moffitt, T. E. (2007). *A review of research on the taxonomy of life-course persistent versus adolescence-limited antisocial behavior.* Cambridge University Press, New York, NY.

Morrow, M. T., Hubbard, J. A., McAuliffe, M. D., Rubin, R. M., & Dearing, K. F. (2006). Childhood aggression, depressive symptoms, and peer rejection: The mediational model revisited. *International Journal of Behavioral Development, 30,* 240- 248.

Morrow, M. T., Hubbard, J. A., Rubin, R. M., & McAuliffe, M. D. (2008). The relation between childhood aggression and depressive symptoms: The unique and joint mediating roles of peer rejection and peer victimization. *Merrill-Palmer Quarterly, 54,* 316-340.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6,* 328-362.

Nocentini, A., Calamai, G., & Menesini, E. (2012). Codevelopment of delinquent and depressive symptoms across adolescence: Time-invariant and time-varying effects of school and social failure. *Journal of Clinical Child & and Adolescent Psychology, 41,* 746-759.

Nolen-Hoeksema, S., & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin, 115,* 424-443.

Obradović, J., Burt, K. B., & Masten, A. S. (2009). Testing a dual cascade model linking competence and symptoms over 20 years from childhood to adulthood. *Journal of Clinical Child & and Adolescent Psychology, 39,* 90-102.

Oldehinkel, A. J., & Bouma, E. M. (2011). Sensitivity to the depressogenic effect of stress and HPA-axis reactivity in adolescence: A review of gender differences. *Neuroscience & Biobehavioral Reviews, 35*(8), 1757-1770.

Ostrov, J. M.., & Crick, N. R. (2005). Current directions in the study of relational aggression during early childhood. *Early Education & Development, 16*, 109-114.

Panak, W. F., & Garber, J. (1992). Role of aggression, rejection, and attributions in the prediction of depression in children. *Development and Psychopathology, 4*, 145-165.

Patterson, G. R., & Capaldi., D. M. (1990). A mediational model for boys' depressed mood. In J. E. Rolf, A. Masten, D. Ciccheti, K. Neuchterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 141-163). Boston: Cambridge University Press.

Patterson, G. R., & Stoolmiller, M. (1991). Replications of a dual failure model for boys' depressed mood. *Journal of Consulting and Clinical Psychology, 59*, 491-498.

Patterson, G. R., & Yoerger, K. (1993). Developmental models for delinquent behavior. In S. Hodgins, S. (Ed.), *Mental disorder and crime* (, pp. 140-172). Thousand Oaks, CA: Sage.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*, 329-335.

Pedersen, S., Vitaro, F., Barker, E. D., & Borge, A. I. H. (2007). The timing of middle-childhood peer rejection and friendship: Linking early behavior to early-adolescent adjustment. *Child Development, 78*, 1037-1051.

Polman, H., de Castro, B. O., Koops, W., van Boxtel, H. W., & Merk, W. W. (2007). A meta-analysis of the distinction between reactive and proactive aggression in children and adolescents. *Journal of Abnormal Child Psychology, 35*, 522-535.

Preacher, K. J., Rucker, D. D., & Hayes, R. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*, 185-227.

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., . . . Liu, J. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32,* 159-171.

Renouf, A., Brendgen, M., Séguin, J. ,R., Vitaro, F., Boivin, M., Dionne, G., . . . Pérusse, D. (2010). Interactive links between theory of mind, peer victimization, and reactive and proactive aggression. *Journal of Abnormal Child Psychology, 38,* 1109-1123.

Rys, G. S., & Bear, G. G. (1997). Relational aggression and peer relations: Gender and developmental issues. *Merrill-Palmer Quarterly, 43,* 87-106.

Salmivalli, C., & Helteenvuori, T. (2007). Reactive, but not proactive aggression predicts victimization among boys. *Aggressive Behavior, 33,* 198-206.

Schmidt, M. E., & Bagwell, C. L. (2007). The protective role of friendships in overtly and relationally victimized boys and girls. *Merrill-Palmer Quarterly, 53,* 439-460.

Schwartz, D., Gorman, A. H., Duong, M. T., & Nakamoto, J. (2008). Peer relations and academic achievement as interacting predictors of depressive symptoms during middle childhood. *Journal of Abnormal Psychology, 117,* 289-299.

Selig, J. P., & Preacher, K. J. (2009). Mediation models for longitudinal data in developmental research. *Research in Human Development, 6,* 144-164.

Smith, R. L., Rose, A. J., & Schwartz-Mette, R. A. (2010). Relational and overt aggression in childhood and adolescence: Clarifying mean-level gender differences and associations with peer acceptance. *Social Development, 19,* 243-269.

Thomas, R., & Zimmer-Gembeck, M. J. (2007). Behavioral outcomes of parent-child interaction therapy and Triple P—Positive Parenting Program: A review and meta-analysis. *Journal of Abnormal Child Psychology, 35,* 475-495.

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort difference on the children's depression inventory: A meta-analysis. *Journal of Abnormal Psychology*, *111*, 578-588.

Underwood, M. K. (2007). Gender and children's friendships: Do girls' and boys' friendships constitute different peer cultures, and what are the trade-offs for development? *Merrill-Palmer Quarterly, 53,* 319-324.

Van der Giessen, D., S., Overbeek, G., Frijns, T., van Lier, P. A. C., Koot, H. M., & Meeus, W. (2013). Co-occurrence of aggressive behavior and depressive symptoms in early adolescence: A longitudinal multi-informant study. *European Review of Applied Psychology, 63,* 193-201.

van Lier, P. A. C., & Koot, H. M. (2010). Developmental cascades of peer relations and symptoms of externalizing and internalizing problems from kindergarten to fourth-grade elementary school. *Development and Psychopathology*, *22*, 569–582.

van Lier, P. A. C., Vitaro, F., Barker, E. D., Brendgen, M., Tremblay, R. E., & Boivin, M. (2012). Peer victimization, poor academic achievement, and the link between childhood externalizing and internalizing problems. *Child Development, 83*, 1775-1788.

Vitaro, F., & Brendgen, M. (2012). Subtypes of aggressive behaviors: Etiologies, development, and consequences. In T. L. Bliesener, A. Beelman, & M. Stemmler (Eds.), *Antisocial behavio and crime: Contributions of developmental and evaluation research to prevention and intervention* (pp. 17-38). Cambridge, MA: Hogrefe.

Vitaro, F., Brendgen, M., & Tremblay, R. E. (2002). Reactively and proactively aggressive children: Antecedent and subsequent characteristics. *Journal of Child Psychology and Psychiatry, 43*, 495-506.

Webster-Stratton, C., Jamila Reid, M., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the incredible years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry, 49,* 471-488.

Wentzel, K. R., & Asher, S. R. (1995). The academic lives of neglected, rejected, popular, and controversial children. *Child Development, 66,* 754-763.

White, B. A., & Kistner, J. A. (2011). Biased self-perceptions, peer rejection, and aggression in children. *Journal of Abnormal Child Psychology, 39,* 645-56.

Appendix A: Study Measures

Proactive and Reactive Aggression (Dodge & Coie, 1987)

	Never	Very Rarely	Some-times	Often	Almost Always
1. When the child has been teased or threatened, he/she gets angry easily and strikes back.	1	2	3	4	5
2. The child always claims that other children are to blame in a fight and feels that they started the trouble.	1	2	3	4	5
3. When someone accidentally hurts the child (such as bumping into him/her), s/he assumes that the peer meant to do it and then reacts with anger/fighting.	1	2	3	4	5
4. The child gets other kids to gang up on somebody that s/he doesn't like.	1	2	3	4	5
5. The child uses physical force (or threatens to use physical force) in order to dominate other kids.	1	2	3	4	5
6. The child threatens or bullies others in order to get his/her own way.	1	2	3	4	5

Withdrawn/Depressed Symptoms (Achenbach & Rescorla, 2001)

	Not True	Somewhat or Sometimes True	Very or Often True
1. There is little he/she enjoys	1	2	3
2. Would rather be alone than with others	1	2	3
3. Refuses to talk	1	2	3
4. Secretive, keeps things to self	1	2	3
5. Too shy or timid	1	2	3
6. Underactive, slow moving, or lacks energy	1	2	3
7. Unhappy, sad, or depressed	1	2	3
8. Withdrawn, doesn't get involved with others	1	2	3

Peer Rejection (Achenbach & Rescorla, 2001)

	Not True	Somewhat or Sometimes True	Very or Often True
1. Doesn't get along with other kids	1	2	3
2. Feels others are out to get him/her	1	2	3
3. Gets teased a lot	1	2	3
4. Not liked by other kids	1	2	3

Academic Performance

	Well Below Average	Below Average	Average	Above Average	Well Above Average
1. How does this child perform academically relative to other students in your class?	1	2	3	4	5
2. When thinking about this student, how would you describe their overall academic performance (reputation based on all their classes)?	1	2	3	4	5
3. What letter grade best reflects this student's academic performance?	F	D	C	B	A

Appendix B: Analyses of Classroom and Cohort Effects

Given the cross-classified structure of the data—with students being nested within (a) *time-invariant study cohorts,* based on their year of entry into the study; and (b) *time-varying classroom teacher IDs*, based on who was their primary classroom teacher at each grade level—supplemental analyses were conducted to examine whether teacher and cohort effects affected the data. Specifically, random-effects ANOVAs, controlling for gender, were estimated to identify group-level mean differences (teacher, cohort, and teacher × cohort), and intraclass correlation coefficients (ICCs) were calculated to estimate effect sizes for the variance accounted for by within-group dependencies in the data. To be conservative, ANOVA results are interpreted without correction for multiple comparisons ($\alpha = .05$), and results are reported for all *p*-values less than .10, even though these are not considered statistically significant (*p*s > .10 are not reported for clarity of presentation).

Results of these analyses are presented in Table B1. As shown, the majority of variables used in the present analyses (i.e., non-shaded cells) did not show statistically significant group differences by teacher ID, cohort, or teacher × cohort interactions (*p*s > .05). However, there were some significant group differences for teacher effects (first grade proactive aggression [$F(6, 10.36) = 15.50$, $p < .001$], first grade reactive aggression [$F(6, 10.15) = 5.55$, $p = .009$], first grade peer rejection [$F(6, 10.18) = 3.74$, $p = .031$], and third grade depressive symptoms [$F(6, 8.80) = 6.35$, $p = .008$]). There were also some teacher × cohort effects (first grade reactive aggression [$F(10, 339) = 2.17$, $p = .020$], and third grade academic performance [$F(9, 362) = 1.92$, $p = .049$]). See B2 tables for group-level descriptive statistics corresponding to teacher main effects, and see B3 tables for teacher × cohort effects. There were no significant main effects of cohort on any study variable.

Similarly, teacher ICCs for study variables ranged from 0 to 0.169, indicating that differences between teachers or classrooms accounted for less than 17% of the variance in any study variable, and about 5% or less for most variables. Cohort ICCs generally fell between 0.00 and 0.01 with the exception of fourth grade academic performance in which cohort effects accounted for 5% of the variance. The higher teacher ICCs (>5%) corresponded precisely with the significant teacher ANOVA effects, occurred primarily among first grade teachers (proactive and reactive aggression and peer rejection) as well as for third grade depressive symptoms.

Taken together, these supplementary analyses reveal that a small portion of the variance for some variables can be accounted for by differences between teachers or classrooms, while other variables show no differences. That is, for some measures and at some grade levels, ratings from the same teacher are likely to be more similar to one another than to ratings from a different teacher. What these analyses do not and cannot reveal is the extent to which these effects are due to teacher rater bias, true differences between classrooms, or both. In the present study, we chose not to model these effects on the rationale that they were of little substantive interest and likely to add a prohibitive degree of complexity to the models estimated. These small ICCs and inconsistent ANOVA results generally support this decision.

Further, it is notable that all of the clustering effects that might be considered non-negligible (e.g., significant ANOVA effects or ICCs > .05) affected only predictor variables, not outcome variables, as specified in all models. Any significant effects from predictor to outcome across grade levels are considered to control for the effects of predictor nestedness through the presumed random assignment of students to classrooms from year to year. Indeed, in most approaches to longitudinally cross-classified data, the goal is to control for the effects of Time T clustering on variables measured at Time T; it is uncommon to account for lagged cross-

classified effects (e.g., of times T-2, T-1 and T on variables measured at time T), and there was no theoretically compelling reason to do this in the present study.

B1 Cohort and Teacher Effects

Table B1. Random effects ANOVAs and intraclass correlation coefficients (ICCs) for the effects of teacher and cohort groups of all study variables.

	Random Effects ANOVAs (F test p-values)				ICCs	
Dependent Variable	Gender (Covariate)	Teacher	Cohort	Teacher × Cohort	Teacher Effects	Cohort Effects
Pro1	.074	<.001	-	-	.169	-
Rea1	<.001	.009	-	.020	.135	-
Aca1	-	-	-	-	-	-
Rej1	.019	.031	-	.073	.091	.003
Dep1	-	<.001	-	-	.185	.004
Pro2	.029	<.001	-	-	.171	.006
Rea2	.004	.069	-	.004	.178	.007
Aca2	-	.011	-	-	.213	.001
Rej2	-	.001	-	-	.132	.002
Dep2	-	<.001	-	-	.357	-
Pro3	-	.060	-	-	.018	-
Rea3	<.001	.026	-	-	.022	-
Aca3	-	.100	-	.049	.054	-
Rej3	-	-	-	-	.020	-
Dep3	-	.008	-	.086	.115	-
Pro4	.058	-	-	.001	-	-
Rea4	<.001	-	-	.055	-	-
Aca4	.074	.081	-	-	.033	.050
Rej4	.015	.087	-	-	.021	-
Dep4	-	.071	-	-	.069	-
Pro5	.079	.001	.037	-	.105	.005
Rea5	.002	.063	.030	-	.028	.020
Aca5	.050	-	-	-	.014	-
Rej5	-	.065	-	-	.020	-
Dep5	-	-	-	-	-	-

Note. P-values are reported for ANOVA terms with significance < .10. Intraclass correlation coefficients are reported where estimable (i.e., random intercept variance > 0). Shaded cells were not relevant to study models, but are reported here for descriptive purposes. See tables below for significant group differences in model variables (nonshaded cells with ANOVA p < .05).

B2. Teacher Effects

Table B2.1. Ratings of 1st grade proactive aggression by teacher ID

Teacher ID	*N*	*M*	*SD*	Min	Max
11	59	2.01	1.08	1.00	4.67
12	57	1.22	0.47	1.00	3.00
13	57	1.22	0.56	1.00	4.33
14	34	1.76	1.01	1.00	4.00
15	57	1.09	0.35	1.00	3.00
16	59	1.15	0.43	1.00	3.33
17	36	1.32	0.71	1.00	4.00
Total	359	1.38	0.76	1.00	4.67

Table B2.2. Ratings of 1st grade reactive aggression by teacher ID

Teacher ID	*N*	*M*	*SD*	Min	Max
11	59	2.86	1.45	1.00	5.00
12	57	2.21	0.93	1.00	4.00
13	57	1.81	1.14	1.00	5.00
14	34	2.67	1.40	1.00	5.00
15	57	1.65	0.88	1.00	5.00
16	59	1.56	0.91	1.00	4.67
17	36	1.83	1.14	1.00	5.00
Total	359	2.07	1.21	1.00	5.00

Table B2.3. Ratings of 1st grade peer rejection by teacher ID

Teacher ID	*N*	*M*	*SD*	Min	Max
11	59	1.27	0.36	1.00	2.25
12	57	1.17	0.22	1.00	2.00
13	57	1.14	0.35	1.00	2.50
14	34	1.38	0.44	1.00	2.50
15	57	1.09	0.21	1.00	2.00
16	59	1.04	0.17	1.00	2.25
17	36	1.15	0.28	1.00	2.00
Total	359	1.17	0.31	1.00	2.50

Table B2.3. Ratings of 3rd grade depressive symptoms by teacher ID

Teacher ID	*N*	*M*	*SD*	Min	Max
33	63	1.25	0.44	1.00	3.00
34	65	1.11	0.24	1.00	2.00
35	62	1.50	0.40	1.00	2.63
36	62	1.10	0.16	1.00	1.63
37	65	1.29	0.46	1.00	2.75
38	41	1.12	0.25	1.00	1.88
39	24	1.21	0.38	1.00	2.50
Total	382	1.23	0.37	1.00	3.00

B3. Teacher × Cohort Interactions

Table B3.1. Ratings of 1st grade reactive aggression by cohort and teacher ID

Teacher ID	Cohort	*N*	*M*	*SD*	Min	Max
11	E	22	2.82	1.77	1.00	5.00
	F	18	3.11	1.27	1.00	5.00
	G	19	2.68	1.22	1.00	5.00
	Total	59	2.86	1.45	1.00	5.00
12	E	20	2.15	0.95	1.00	4.00
	F	18	2.28	1.09	1.00	4.00
	G	19	2.21	0.78	1.00	3.67
	Total	57	2.21	0.93	1.00	4.00
13	E	23	1.39	0.64	1.00	3.00
	F	17	2.35	1.46	1.00	5.00
	G	17	1.84	1.14	1.00	5.00
	Total	57	1.81	1.14	1.00	5.00
14	F	17	3.02	1.58	1.00	5.00
	G	17	2.31	1.13	1.00	4.33
	Total	34	2.67	1.40	1.00	5.00
15	E	23	1.77	0.92	1.00	5.00
	F	17	1.35	0.68	1.00	3.00
	G	17	1.80	0.97	1.00	4.00
	Total	57	1.65	0.88	1.00	5.00
16	E	23	1.65	1.05	1.00	4.67
	F	18	1.69	0.86	1.00	3.33
	G	18	1.33	0.77	1.00	3.00
	Total	59	1.56	0.91	1.00	4.67
17	F	17	1.35	0.67	1.00	3.33
	G	19	2.26	1.31	1.00	5.00
	Total	36	1.83	1.14	1.00	5.00
Total	E	111	1.94	1.22	1.00	5.00
	F	122	2.17	1.30	1.00	5.00
	G	126	2.07	1.12	1.00	5.00
	Total	359	2.07	1.21	1.00	5.00

Table B3.2. Ratings of 3rd grade academic performance by cohort and teacher ID

Teacher ID	Cohort	*N*	*M*	*SD*	Min	Max
33	C	21	2.68	1.32	1.00	5.00
	D	20	3.90	1.15	2.33	5.00
	E	22	3.02	1.56	1.00	5.00
	Total	63	3.19	1.43	1.00	5.00
34	C	20	3.97	0.79	2.33	5.00
	D	21	3.52	1.08	1.67	5.00
	E	24	3.82	0.87	1.67	5.00
	Total	65	3.77	0.93	1.67	5.00
35	C	18	3.37	0.92	1.33	4.67
	D	21	3.19	0.94	1.00	5.00
	E	23	3.30	1.15	1.33	5.00
	Total	62	3.28	1.00	1.00	5.00
36	C	19	3.75	1.26	1.33	5.00
	D	20	4.15	1.27	1.00	5.00
	E	23	4.06	1.27	1.33	5.00
	Total	62	3.99	1.25	1.00	5.00
37	C	19	3.23	0.78	1.67	5.00
	D	21	3.19	0.36	2.33	3.33
	E	24	3.24	0.43	2.33	4.33
	Total	64	3.22	0.53	1.67	5.00
38	C	20	3.47	0.93	1.67	5.00
	D	21	3.62	0.97	1.67	5.00
	Total	41	3.54	0.94	1.67	5.00
39	E	24	3.49	1.11	1.67	5.00
	Total	24	3.49	1.11	1.67	5.00
Total	C	117	3.40	1.09	1.00	5.00
	D	124	3.59	1.04	1.00	5.00
	E	140	3.49	1.15	1.00	5.00
	Total	381	3.50	1.09	1.00	5.00