THE ORIGINS OF ANTISOCIAL BEHAVIOR: INFLUENCES OF GENES, DEVELOPMENT, AND CONTEXT

By

Sarah Lynn Carroll

A DISSERTATION

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

Psychology – Doctor of Philosophy

2024

ABSTRACT

Developmental trajectories of youth antisocial behavior (ASB; i.e., physical aggression and rule violations) unfold via dynamic interplay between individual predispositions and the contexts in which their development is embedded. Indeed, prior work has implicated both genetic influences and the broader environmental context (i.e., family, peers, and neighborhoods) in the emergence of youth behavioral problems, both cross-sectionally and over time. Nevertheless, much remains unknown about the developmental origins of ASB. Extant longitudinal studies examining its genetic and environmental etiology have done so across relatively brief windows of development, and nearly all longitudinal phenotypic studies have relied on either person-centered or variable-centered approaches but not both, with key downstream consequences for our understanding of the factors that give rise to the development of ASB. The present studies addressed these gaps in the literature by examining trajectories of ASB across early development via a series of behavioral genetic, variable-centered, and person-centered analyses, with familial, peer, and neighborhood characteristics considered as moderators. Participants were drawn from the Twin Study of Behavioral and Emotional Development in Children (TBED-C; *N* = 1,030 twin pairs) and MTwiNS projects within the Michigan State University Twin Registry. Both the TBED-C and its longitudinal follow-up, MTwiNS, were enriched for neighborhood disadvantage. Study 1 made use of a series of longitudinal growth curve models and classical twin analyses to examine the genetic and environmental origins of ASB from the preschool years through to emerging adulthood, with neighborhood disadvantage considered as a phenotypic predictor. Results indicated a mean-level decline in ASB that was shaped by both genetic influences and nonshared (i.e., person-specific) environmental factors. Exposure to neighborhood disadvantage predicted higher levels of ASB at baseline but a somewhat more rapid age-related decline. Studies 2 and 3 built on these findings by disambiguating physical aggression (AGG) and non-aggressive rule-breaking (RB), respectively, and by incorporating person-centered statistical methods as well as variable-centered methods. In Study 2, the development of AGG was examined from middle childhood through to emerging adulthood via growth curve modeling and mixture modeling. Both sets of analyses identified age-related declines in AGG, but only the person-centered models elucidated the factors (i.e., low parent-child conflict and familial affluence) that interrupted trajectories of elevated AGG. Likewise, in Study 3, variable-centered analyses were integral to modeling the mean-level increase in RB for the full sample, whereas the person-centered models implicated peer delinquency as well as family and neighborhood socioeconomic disadvantage as predictors of persistent and escalating RB trajectories. Altogether, these findings underscore the non-interchangeable contributions of the family, neighborhood, and peer contexts to trajectories of ASB across nearly the entire early developmental period. Moreover, integration of findings from variable-centered and person-centered statistical approaches demonstrated the potential to illuminate risk factors predicting persistent behavioral problems and protective factors predicting recovery, which may have important methodological implications for future studies of developmental psychopathology.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without Dr. Alex Burt's guidance, expertise, and enthusiasm for twin research and Dr. Kelly Klump's, Dr. Alytia Levendosky's, and Dr. Amber Pearson's insightful feedback as part of my committee. I am deeply grateful to Dr. Burt, my committee chair and advisor, for always challenging me to become a better researcher and for believing that it was, in fact, possible for me to integrate behavior genetic and developmental methods. I am also grateful for the support of all of the clinical faculty, particularly Dr. Levendosky, whose kind words of encouragement throughout grad school have always lifted me up. I would also like to thank my funding sources, including the National Institute of Child Health and Human Development and the Michigan State University Graduate School, as well as the research assistants and staff in the Burt Lab for their substantial time and effort in collecting data.

My journey in clinical psychology began in Dr. Eric Turkheimer's lab at the University of Virginia. I am thankful for Dr. Turkheimer's continuous support since my first days as a research assistant, and particularly for his generosity in helping me with ridiculously complicated statistical analyses every spring. Getting to see him at the Behavior Genetics Association's meeting each year is always a conference highlight. I am also immeasurably grateful to Erinn Falls and Stephanie Rudlaff, whose support has made such a difference in my life both personally and professionally.

I am also thankful for my parents, brother, and friends, who have supported me in my work toward becoming a clinical psychologist. I couldn't have imagined a better friend than Megan Mikhail to walk with me through grad school, and I am so very excited to watch (the future) Dr. Mikhail's career continue to unfold. In addition, Jack Baker has been a continuous source of support and affirmation, particularly in my clinical training; he is one of several clinicians who I admire and seek to emulate. I am also thankful for Maddy Baker, who has been a dear friend for more than 20 years and will be for the rest of my life, and Ben Vogel, whose

humor and warmth are such a delight. Throughout grad school, I have also fondly remembered my wonderful grandmother, Ruby, whose nursing career and deep compassion for others continue to inspire me.

I have found so much meaning and joy in being part of the Eastminster Presbyterian Church community over the past six years. Thank you to the members of Eastminster for welcoming, encouraging, and loving me since my first week in East Lansing. To the extent that I've been able to embody that same welcoming spirit in the therapy room, I'm a better clinician for having known the people there. I am also grateful to the clients with whom I've gotten to work, some on a long-term basis and others briefly. Being able to walk with them for any length of time in their efforts to become happier, healthier, and more self-confident has been a privilege and a joy. I am grateful to my clinical supervisors, particularly Dr. Natalie Moser, Dr. Jason Stentoumis, and Dr. Carol Laub, for bolstering my confidence and providing an outstanding example of what ethical, compassionate mental healthcare looks like.

Lastly, I am thankful for my precious rescue cats, Francie and Laurie, who have been a continuous source of cuddles and comic relief.

TABLE OF CONTENTS

GENERAL INTRODUCTION

Youth antisocial behavior (ASB) encompasses a broad range of behaviors that violate societal norms and/or the basic rights of others (American Psychiatric Association, 2013). Children who demonstrate high levels of behavior problems often do so throughout early development, culminating in a series of poor outcomes by adulthood (e.g., poor mental and physical health, financial difficulties, incarceration; Odgers et al., 2008). Despite this high level of rank-order stability, however, ASB is far from a static trait. Indeed, a large body of literature has found the absolute frequency of ASB and its specific behavioral manifestations to vary dramatically with age. Specifically, engagement in physical aggression (AGG; e.g., hitting, kicking) declines throughout childhood and adolescence for most youth, whereas non-aggressive rule-breaking (RB; e.g., lying, stealing) typically peaks in frequency during adolescence (Burt, 2012; Moffitt, 1993, 2003). Moreover, there are pronounced individual differences in the development of ASB, with youth differing in their engagement at baseline as well as in the magnitude and direction of their change over time (e.g., Doornward et al., 2012; NICHD Early Child Care Research Network & Arsenio, 2004).

Individual differences in the frequency and specific manifestation of ASB have been linked to numerous factors. For instance, biological sex consistently predicts engagement in ASB, with males demonstrating markedly higher rates of physical AGG, and somewhat higher rates of RB, across the lifespan relative to females (Moffitt, 2003). Differences in the frequency of ASB are also observed across families and across societies (Rescorla et al., 2007), reflecting genetic influences as well as the effects of the rearing environment (Burt, 2022). The neighborhood and peer contexts are also critical predictors, such that children exposed to concentrated poverty and/or delinquent peers exhibit far more behavioral problems on average than other youth (Leventhal & Brooks-Gunn, 2000; Mann et al., 2015). Findings from longitudinal studies suggest that discrepancies in ASB engagement across neighborhoods persist throughout adolescence and into early adulthood (e.g., Lacourse, Dupéré, & Loeber, 2008; Maldonado-Molina, Reingle, et al., 2010; Spano et al., 2010).

Although extant longitudinal research has shed light on trajectories of ASB, much remains unknown about its developmental origins. For instance, a small body of research has leveraged both longitudinal and classical twin methods to examine the genetic and environmental origins of ASB (e.g., Eley et al., 2003; Burt, McGue, Carter, & Iacono, 2007), but most studies have done so across a relatively brief window of development (e.g., middle childhood through early adolescence). As a result, it is unclear how genetic predispositions and environmental risk/protective factors differentially affect stability and change in ASB across multiple developmental stages. The genetic factors contributing to early childhood ASB, for example, could be the same ones that underlie stability during late adolescence, which would indicate that the rank-order stability in ASB observed across early development is largely genetic in origin. Alternatively, and consistent with research identifying age-related increases in the heritability of ASB (Rhee & Waldman, 2002), novel genetic influences may come into play during adolescence and contribute to systematic change. It is also unclear whether environmental influences on ASB are largely transient and unsystematic, or whether these factors systematically influence behavioral trajectories. In short, because no genetically informative study to date has examined continuity and change in ASB across all of early development, it is unclear to what extent genetic and environmental influences, respectively, are continuous from early childhood through emerging adulthood.

Extant phenotypic studies have also yielded a somewhat limited picture of development. Specifically, almost all longitudinal studies rely on only one of two statistical approaches: variable-centered or person-centered. Variable-centered analyses use growth curve models to quantify the mean change(s) over time in the outcome of interest across the entire sample (McArdle & Epstein, 1987; see Figure 1.0a), whereas person-centered approaches extract subgroups exhibiting distinct trajectories over time (Nagin, 1999; see Figure 1.0b). Each of these approaches necessarily provides a *partial* picture of development. In variable-centered studies, for instance, the growth curve representing average change may not capture the heterogeneity observed in most samples. The primary strength of person-centered methods lies in their ability to identify subgroups of youth following trajectories of either persistent engagement or recovery, as well as elucidate the risk and protective factors, respectively, that distinguish these youth from their peers. Nevertheless, person-centered studies of ASB are critiqued for subdividing a continuous trait into qualitatively distinct groups, with some researchers arguing that ASB is best-represented using variable-centered models (Walters, 2011; Walters & Ruscio, 2013). Given the unique strengths and limitations of each method, the use of both is likely to yield a far more complete picture of development than either could provide by itself.

In sum, ASB is a complex trait, influenced by a myriad of genetic and environmental factors that may or may not be continuous over time. Moreover, the frequency of engagement varies considerably within persons over time as well as between persons at any given timepoint. To represent the development of ASB in a way that captures this complexity, longitudinal studies must leverage a wide range of statistical methods that disentangle genetic and environmental influences and model continuity and change not only across individuals but also across trajectory groups. The present studies sought to do just this by examining trajectories of ASB from childhood through to emerging adulthood via a series of variable-centered, person-centered, and classical twin analyses. As detailed below, we first elucidated the genetic and environmental origins of ASB trajectories in general. We then examined the phenotypic development of AGG and RB, respectively, through two complementary sets of longitudinal analyses, which each yield distinct information about the ways in which family and neighborhood characteristics shape behavioral development. Thus, the studies included in this dissertation collectively aimed to advance our understanding of the dynamic developmental processes that give rise to youth externalizing trajectories.

To accomplish this, the present studies made use of longitudinal data from the Twin

Study of Behavioral and Emotional Development in Children (TBED-C), a sample within the population-based Michigan State University Twin Registry and the only twin sample in the world specifically enriched for neighborhood disadvantage (MSUTR; Burt & Klump, 2019). Behavioral and emotional data relevant to the present studies were collected at up to five time points (between ages three and 22 years). Thus, this set of studies addressed the primary limitations of prior longitudinal work by examining engagement in ASB across multiple developmental stages and integrating results from two distinct statistical approaches. All studies examined the moderating role of neighborhood disadvantage on trajectories of continuity and trajectories of change in youth ASB.

Overview of the three studies

Study 1 examined the origins of stability and change in ASB from preschool through to emerging adulthood, with biological sex and neighborhood disadvantage considered as phenotypic predictors. Extant longitudinal twin studies have consistently found genetic influences to underlie rank-order stability in ASB, and nonshared environmental influences (e.g., differential parenting, peer groups) to contribute to change (e.g., Bartels et al., 2004; Burt et al., 2007; Eley et al., 2003). The role of the shared environment (e.g., socioeconomic status, family mores) has appeared to be dependent on age, with studies in childhood samples reporting shared environmental influences to contribute to stability (e.g., van Beijsterveldt et al., 2003) while studies in older samples have found their effects to be negligible (e.g., Burt et al., 2007). However, as discussed above, these studies all assessed continuity and change in ASB over relatively brief time periods. Because no prior study has examined the etiology of ASB across all of early development, it is unclear to what extent genetic factors underlie stability in ASB from the preschool years through emerging adulthood, nor is it clear how shared and nonshared environmental factors differentially affect stability and change across multiple developmental stages. The present study thus made use of variable-centered multilevel growth curve modeling to examine the development of ASB from ages three to 22 years in participants from the TBED-

C (*N* = 1,030 twin pairs) assessed up to five times. Biological sex and neighborhood disadvantage, defined via a composite index of area deprivation (see Table 1.0), were examined as predictors of the phenotypic intercept and slope. Classical twin modeling was then used to quantify the genetic, shared environmental, and nonshared environmental contributions to participants' baseline levels of ASB (i.e., intercepts) and rates of change (i.e., slopes).

Studies 2 and 3 built on Study 1 by 1) modeling behavioral trajectories via two distinct statistical approaches, and 2) distinguishing between physically aggressive and non-aggressive dimensions of ASB, respectively. As discussed earlier, variable-centered and person-centered approaches are rarely integrated, with no extant empirical study considering the ways in which their respective results may complement one another. The results from extant variable-centered and person-centered literatures were conceptually integrated in a recent review of longitudinal studies of ASB (Carroll, Mikhail, & Burt, 2023), but no empirical study to date has integrated the two approaches. Study 2 began to fill this gap in the literature by leveraging both approaches to examine desistance from AGG between childhood and emerging adulthood. We focused specifically on behavioral trajectories across middle childhood and adolescence, two key developmental periods characterized by significant changes in the frequency of engagement in ASB (Moffitt, 1993). Biological sex, family characteristics (i.e., household income, parent-child conflict, and parent-child nurturance), and neighborhood disadvantage were each evaluated as phenotypic moderators of AGG trajectories, with a focus on the *protective factors* predicting desistance. Study 3 leveraged the same modeling approaches to examine the development of RB, which follows a typical trajectory that is markedly different from that of AGG (Burt, 2012). Specifically, Study 3 sought to elucidate the *risk factors* predicting persistent and escalating engagement in RB, a pattern often observed during adolescence (e.g., Yan et al., 2021). Biological sex, household income, neighborhood disadvantage, and peer delinquency were examined as moderators of RB trajectories.

Altogether, by conducting genetically informative analyses of ASB across multiple developmental periods and integrating findings from two distinct statistical approaches, the present studies substantially advance our understanding of the patterns of desistance, persistence, and escalation that characterize the development of ASB across time and persons. Moreover, the demonstrated utility of variable-centered and person-centered approaches together may have important implications for future work examining trajectories of psychopathology.

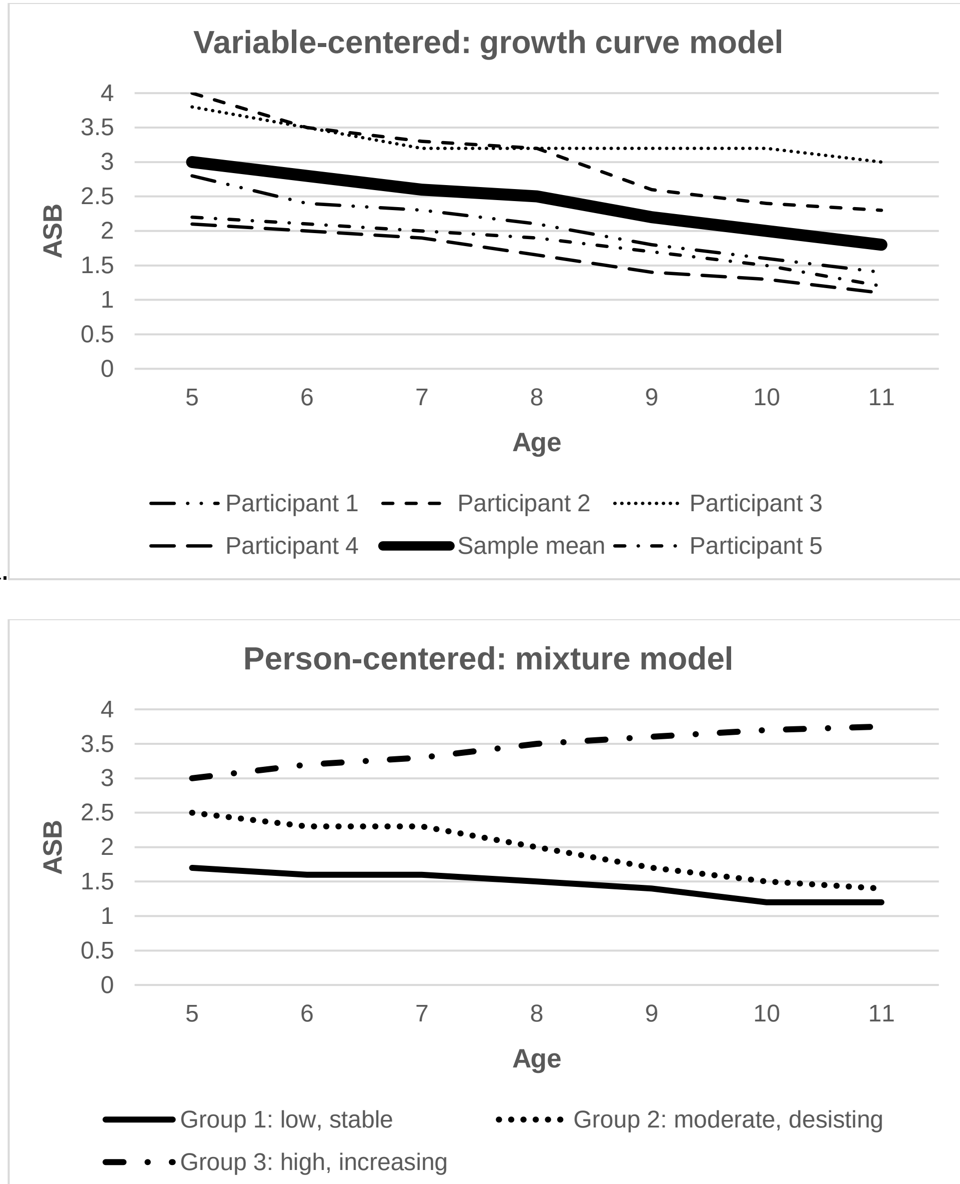
Tables

Table 1.0. *Measures included in Area Deprivation Index*

Measure
1. Percent of population aged 25 and older with <9 years of education
2. Percent of population aged 25 and older with at least a high school diploma
3. Percent of population aged 16 and older in white-collar occupations
4. Median family income
5. Income disparity (ratio of households with <$10,000 income to households with ≥$50,000 income)
6. Median home value
7. Median gross rent
8. Median monthly mortgage
9. Percent of housing units owned by occupiers
10. Percent of population aged 16 and older who are unemployed
11. Percent of families below poverty level
12. Percent of population below 150% of the poverty threshold
13. Percent of households with children under age 18 headed by a single parent
14. Percent of households without a motor vehicle
15. Percent of households without a telephone
16. Percent of occupied housing units without complete plumbing
17. Percent of households with more than 1 person per room

Note: Participating families' ADI scores were determined by the level of deprivation in their Census tract (Study 1) and block group (Studies 2 and 3) based on all indices listed above. For additional details, see Singh (2003) and Kind & Buckingham (2018).

Figures



a.

b.

Figure 1.0. *Variable-centered (a) and person-centered (b) approaches to modeling the development of ASB across middle childhood in a hypothetical sample.* The variable-centered approach identified a mean-level trend across the entire sample, whereas the person-centered analysis extracted three distinct trajectory groups.

CHAPTER 1: CONTINUITY AND CHANGE IN THE GENETIC AND ENVIRONMENTAL ETIOLOGY OF YOUTH ANTISOCIAL BEHAVIOR

Abstract

Background: Trajectories of youth antisocial behavior (ASB) are characterized by both continuity and change. Twin studies have further indicated that genetic factors underlie continuity, while environmental exposures unique to each child in a given family underlie change. However, most behavioral genetic studies have examined continuity and change during relatively brief windows of development (e.g., during childhood but not into adolescence). It is unclear whether these findings would persist when ASB trajectories are examined across multiple stages of early development (i.e., from early childhood into emerging adulthood). **Methods:** Our study sought to fill this gap by examining participants assessed up to five times between the ages of 3 and 22 years using an accelerated longitudinal design in the Michigan State University Twin Registry (MSUTR). We specifically examined the etiologies of stability and change via growth curve modeling and a series of univariate and bivariate twin analyses. **Results:** While participants exhibited moderate-to-high rank-order stability, mean levels of ASB decreased linearly with age. Genetic and nonshared environmental influences that were present in early childhood also contributed to both stability and change across development, while shared environmental contributions were negligible. In addition, genetic and nonshared environmental influences that were not yet present at the initial assessment contributed to change over time. **Conclusions:** Although ASB tended to decrease in frequency with age, participants who engaged in high levels of ASB during childhood generally continued to do so throughout development. Moreover, the genetic and nonshared environmental contributions to ASB early in development also shaped the magnitude of the decrease with age.

Introduction

Childhood antisocial behavior (ASB) predicts a myriad of poor outcomes in adolescence and young adulthood, such as substance use, poor physical health, and internalizing pathology, as well as continued engagement in delinquent activities (e.g., Odgers et al., 2008). In children, ASB is characterized by persistent aggression, deceitfulness, property destruction, and/or rule violations (American Psychiatric Association, 2013). One of the defining features of ASB is its relatively high level of rank-order stability, with the same individuals typically exhibiting the highest levels of delinquent behavior across development. Despite this stability, mean levels of ASB decline throughout the first twenty or so years of life (e.g., Monahan et al., 2009), and there is considerable individual variation in the magnitude of this decline (Burt, 2012; Martino et al., 2008). In short, extant research has clearly indicated that youth ASB trajectories are characterized by both stability and change, and that these patterns vary from person to person.

To date, however, it is less clear what etiologic mechanisms underlie individual differences in these patterns of continuity and change. There are several competing a priori possibilities. The relatively high rank-order stability of ASB could be due to continuity in underlying genetic influences, while change over time could stem from specific environmental exposures (e.g., environmental risk factors could predict escalating behavior problems while environmental protective factors predict desistance). Alternatively, genetic contributions could change over time as different genes become (de)activated, while environmental factors could contribute to stability. One method for evaluating these competing hypotheses is the twin design, which compares monozygotic (MZ) and dizygotic (DZ) twins to disambiguate the genetic and environmental contributions to a given phenotype. By examining these contributions across multiple timepoints, twin researchers can clarify the origins of continuity and change.

Prior longitudinal twin studies have begun to evaluate these possibilities. These studies have consistently implicated genetic factors as a major source of rank-order stability in ASB (Bartels et al., 2004; Burt et al., 2007; Eley et al., 2003; Lacourse et al., 2014; Pingault et al., 2015; Porsch et al., 2016; van Beijsterveldt et al., 2003). As an example, Burt et al. (2007) used latent growth curve modeling to examine the etiologic trajectory of ASB from late adolescence through early adulthood (approximately ages 17-25) in 626 twin pairs from the Minnesota Twin Family Study and found the same genetic factors to be present over time. These factors explained a moderate-to-large proportion of the variance in ASB at each timepoint and were largely responsible for trait stability. Nonshared environmental influences (or experiences that serve to differentiate children raised in the same family; e.g., peer groups) were found to underlie change over time. Shared environmental influences (experiences common to children raised in the same family; e.g., similar parenting) did not contribute to ASB at baseline or over time, consistent with research indicating that shared environmental influences on ASB become less salient (and nonshared more salient) with age (e.g., Tuvblad et al., 2011). Child and adolescent twin studies using liability threshold analyses, simplex modeling, or Cholesky decomposition modeling have reported somewhat similar results, finding that nonshared environmental factors largely contribute to change over time, while genetic influences contribute to stability (Bartels et al., 2004; Eley et al., 2003; van Beijsterveldt et al., 2003). However, these studies have also found evidence that shared environmental influences contribute to stability over time. In other words, shared environmental influences have been found to impact ASB development in younger samples, but their effects appear to be negligible by late adolescence.

Despite these consistencies in results, several questions remain unanswered. First, it is unclear whether the genetic factors contributing to adolescent and young adult ASB are the same as those contributing to child ASB. Although the studies discussed above found largely continuous genetic effects, most assessed continuity and change during a relatively brief window of development: early childhood through pre-adolescence, or middle childhood through early adolescence, or late adolescence through early adulthood. The most comprehensive etiologic study of ASB development (Pingault et al., 2015) began in early childhood (twins were assessed beginning at age 4 through 16 years) but did not assess participants in late adolescence or emerging adulthood, two key developmental periods in the transition to adult social and occupational roles (Alink & Egeland, 2013). Indeed, no study to date has examined the etiologies of continuity and change in ASB across all of early development (i.e., the first twenty or so years of life). As such, we do not know to what extent genetic factors underlie stability in ASB from the preschool years through emerging adulthood, nor is it clear how shared and nonshared environmental factors differentially affect stability and change throughout this period.

The latter uncertainty is particularly important, given that the shared environment has been identified as an important etiologic source of stability in ASB during childhood and early adolescence (Bartels et al., 2004; Eley et al., 2003; van Beijsterveldt et al., 2003), but does not appear to affect the trajectory of ASB during late adolescence (Burt et al., 2007). By contrast, nonshared environmental influences appear to be transient and idiosyncratic prior to adulthood, with nonshared environmental correlations for positive and negative affect and interpersonal warmth each decreasing monotonically in a matter of minutes or days (Burt et al., 2015). At some point during adolescence, however, these influences appear to become more enduring. Although a few studies have reported this increased stability of the non-shared environment as early as age 7 (e.g., van Beijsterveldt et al., 2003), most studies place it sometime in late adolescence (e.g., Hopwood et al., 2011). As such, results from studies of children identifying the nonshared environment as a source of change may not persist to older samples. In short, much is unknown about the developmental origins of ASB because no study of its etiologic trajectory has spanned early childhood through emerging adulthood.

The aim of the present study was to address these gaps in the literature by examining the origins of stability and change in youth ASB from preschool through to emerging adulthood. We used up to five waves of data from the Michigan State University Twin Registry (MSUTR; Burt & Klump, 2019) collected across ages 3 to 22 years using an accelerated longitudinal design. We applied multilevel growth curve modeling, in which measurements were nested within participants, to estimate participants' baseline levels of ASB (i.e., intercepts) and rates of change with age (i.e., slopes). We subsequently used classical twin modeling to quantify the genetic, shared environmental, and nonshared environmental influences on the intercept and slope, respectively. Based on prior research, we hypothesized that genetic factors would underlie stability and nonshared environmental factors would underlie change. Given that shared environmental effects have been found to decrease with age, we did not expect them to significantly impact participants' trajectories into emerging adulthood.

Methods

Participants

Participants were drawn from the Twin Study of Behavioral and Emotional Development in Children (TBED-C), a sample within the population-based Michigan State University Twin Registry (MSUTR; Burt & Klump, 2019). The TBED-C includes both a population-based subsample (n=528 families) and an independent 'at risk' subsample of twin families residing in impoverished Census tracts (n=502 families). When combined, the overall sample thus comprised 1,030 twin pairs: 224 MZ male pairs, 211 DZ male pairs, 202 MZ female pairs, 207 DZ female pairs, and 186 DZ opposite-sex pairs. Mean household income at the middle childhood assessment was $76,329 (SD=$45,650) in the population-based sample and $55,652 (SD=$31,088) in the at-risk sample. Other recruitment details are reported at length in prior publications (e.g., Burt & Klump, 2019). Families across the two samples collectively identified as White (non-Latinx): 81%, Black: 10%, Latinx: 1%, Asian: 1%, Indigenous: 1%, and multiracial: 6%. These proportions are largely consistent with those for the population of the State of Michigan, based on data from the U.S. Census Bureau (http://www.Census.gov/) (e.g., White: 79%, Black: 14%). For all studies, parents provided informed consent, children provided informed assent, and families were compensated for their time.

Behavioral and emotional data relevant to the current study were collected at as many as five time points. All 1,030 twin families were assessed once in middle childhood (ages 6-11) as part of the TBED-C. Those TBED-C twins residing in modestly-to-severely disadvantaged neighborhoods are currently being reassessed in-person as adolescents up to two times, 18-months apart, through the Michigan Twin Neurogenetics Study (MTwiNS). The first of the MTwiNS assessments was conducted approximately 4-6 years after participation in TBED-C (ages 7-19; currently available for 354 families), while the second adolescent assessment was 5-7 years after participation in TBED-C (ages 10-19; currently available for 188 families). TBED-C families with twins between ages 11 and 22 were also recently recruited for an online assessment of youth psychopathology (N=637 families completed the online assessment). Finally, we were also able to link to data collected on TBED-C families as part of the population-based Michigan Twins Project (MTP), an ongoing study of approximately 12,000 Michigan-born child and adolescent twin pairs (93.3% of TBED-C families were recruited out of the MTP). See Table 1.1 for additional details about the sample at each age, including sample sizes at each assessment.

Notably, the MTP assessments were not completed in the same order or at the same ages across participating TBED-C families. For example, while most families were first assessed as part of the MTP (73.9%), others were first assessed as part of TBED-C and were assessed only later as part of the MTP (26.1%). Follow-up MTP assessments are also on-going. A portion of TBED-C families (N = 637; 61.8%; 56 of these were assessed twice) were re-assessed approximately 5-8 years after their original participation in the MTP. To account for these irregularities in the ordering of data collection across the TBED-C/MTwiNS and the MTP, data were organized <u>chronologically by age</u> for each participating twin. As such, each assessment wave in the current study includes MTP and either TBED-C or MTwiNS assessments (see Table 1.2). A total of 677 pairs (66% of the sample) completed at least three of the assessments, while 96% (N=989) completed at least two. The highest number of assessments completed by any single participant was five.

Recruitment

Because birth records are confidential in Michigan, we collaborated with the Michigan Department of Health and Human Services (MDHHS; formerly known as the Michigan Department of Community Health) to recruit families for all MSUTR twin studies (including the MTP and all waves of the TBED-C/MTwiNS). The MDHHS is the agency in charge of all vital records in the State of Michigan and thus has direct access to individual SSNs, full names, and birth dates. The MDHHS identifies twin pairs residing in Lower Michigan who meet age criteria for a given study and whose addresses or parents' addresses (for twins who are minors) can be located either using driver's license information obtained from the State of Michigan or the proprietary search engine used by police. Twins indicating interest in participation via pre-stamped postcards or e-mails/calls to the MSUTR project office are then contacted by study staff to determine study eligibility and to schedule their assessments.

Four recruitment mailings were used to ensure optimal twin participation. Overall, response rates across studies (56-85%) are on par with or better than those of other twin registries that use similar types of anonymous recruitment mailings and have thus far yielded largely representative samples. Families of the naturally-conceived twins in the large-scale MTP, for example, closely resemble families across the State of Michigan (Burt & Klump, 2013). The proportion of MTP families that identify as White, non-Hispanic (81.0%) is very similar to the 80.2% indicated in state-wide Census data. Mean family incomes are also quite comparable ($75,940 in the MTP versus $73,373 in the Census), as are the proportion of families with graduate or professional degrees (10.3% in the MTP versus 9.6% in the Census).

Because 90+% of TBED-C families were recruited out of the MTP, we were able to use the MTP data to compare families who chose to participate in TBED-C with those who were recruited but did not participate. TBED-C families were generally representative of recruited but non-participating families. As compared to non-participating twins, participating twins reported similar levels of conduct problems, emotional symptoms, and hyperactivity (*d* ranged from -.08

to .01 in the population-based sample and .01 to .09 in the at-risk sample; all ns). Participating families also did not differ from non-participating families in paternal felony convictions (d = -.01 and .13 for the population-based and the at-risk samples, respectively), rate of single parent homes (d = .10 and -.01 for the population-based and the at-risk samples, respectively), paternal years of education (both $d \leq .12$), or maternal and paternal alcohol problems (d ranged from .03 to .05 across the two samples). However, participating mothers in both samples reported slightly more years of education (d = .17 and .26, both $p < .05$) than non-participating mothers. Maternal felony convictions differed across participating and non-participating families in the population-based sample (d = -.20; $p < .05$) but not in the at-risk sample (d = .02). In short, our recruitment procedures thus appear to yield samples that are representative of both recruited families and the general population of the State of Michigan.

Procedure

Some of the assessments, specifically the MTP assessments and the "online" assessment, were completed remotely by the twins' primary caregiver, nearly always their mother. The TBED-C and MTwiNS assessments were completed in-person. For TBED-C, twins were assessed either at our East Lansing-based laboratories or at the family's home. Questionnaires did not vary across the laboratory-based and family home assessments. For MTwiNS, the twins and their parent(s) completed an in-person assessment lasting 4-8 hours at either the East Lansing or Ann Arbor-based laboratories.

Measures

Youth Antisocial Behavior

Youth ASB was assessed via maternal report at all ages. At the MTP assessments, participating twins' mothers completed the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001), a 25-item measure in which parents rate the extent to which a series of statements describe the child's behavior over the past six months using a three-point scale (0=not true to 2=certainly true). For these analyses, we focused on the Conduct Problems subscale (5 items: hot temper, obedient (reverse-scored), fights, lies or cheats, steals; α=.60, .63, and .66 at MTP assessments 1, 2, and 3, respectively). Psychometric studies have found the SDQ to have satisfactory test-retest reliability (r>.85 for the Conduct Problems subscale) and to be highly correlated with other parent-report measures, including the Child Behavior Checklist (CBCL) (e.g., Muris et al., 2003). In addition, studies in samples spanning childhood and adolescence support the use of the parent-report SDQ as a screening measure that adequately distinguishes between community and clinical populations across age groups (Becker et al., 2004; He et al., 2013).

At the in-person TBED-C and MTwiNS assessments, the twins' mothers completed the CBCL (Achenbach & Rescorla, 2001), rating the extent to which a series of statements described the child's behavior during the past six months on a three-point scale (0=never to 2=often/mostly true). To maximize comparability with the SDQ, we constructed a scale using 5 items on the CBCL that were analogous to those on the SDQ: hot temper, disobedient at home, fights, lies or cheats, steals from home (α=.66 in the TBED-C and .69 and .60, respectively, at the two MTwiNS assessments). While these items screen for behaviors that may manifest differently at different ages (e.g., lying, hot temper), all are relatively common across early development. As such, we believe that they adequately assess ASB across the broad age range included in our sample.

Zygosity

Zygosity was established using physical similarity questionnaires administered to the twins' primary caregiver (Peeters et al., 1998). On average, the physical similarity questionnaires used by the MSUTR have accuracy rates of at least 95% when compared to DNA.

Disadvantage

Socioeconomic disadvantage was assessed using the Area Deprivation Index (ADI), a composite measure comprising 17 indices of Census-tract disadvantage (e.g., poverty rate, income disparity). We recreated Kind & Buckingham's index of disadvantage in our sample, as assessed via Census data collected from 2008 to 2012. The measures were weighted according to the factor loadings identified by Kind & Buckingham (2018), and weighted variables were summed to create a deprivation index score for each Census tract. Families were assigned a percentile indicating the level of deprivation in their Census tract relative to that of all Census tracts in Michigan. The mean ADI was 42.51 (SD=26.17) and ranged from 1 to 100.

Data Analyses

Phenotypic Analyses

All analyses were conducted using M*plus* 8.0 (Muthén & Muthén, 1998-2019). To examine phenotypic changes in ASB over time, we used a three-level growth curve model in which occasions of measurement (Level 1) were nested within participants (Level 2) who were nested within families (Level 3). These models capture both the average rate of change in ASB over the course of the study, as well as individual variability in change via random intercept and slope terms. Age was used as the index of time in these models and was centered at three years old, the youngest age in our sample. The intercept can thus be interpreted as the level of ASB at age 3.

We initially estimated an unconditional growth model with a random linear slope, which allowed for interindividual variation in the rate of change over time. Models with non-linear slopes encountered serious convergence difficulties. Moreover, prior research consistently indicates that mean levels of ASB decline steadily throughout development (e.g., Monahan et al., 2009). We subsequently estimated a conditional growth model. In this model, questionnaire type (SDQ or CBCL) was added as a time-varying covariate on level 1 with a random slope. We included biological sex, ethnicity, and ADI as covariates of the random intercept and slope on level 2.

Full-information maximum likelihood estimation was used to account for missing data, as prior simulations have shown it to be robust to at least 50% missing data (Enders & Bandalos, 2001). Moreover, accelerated longitudinal designs such as ours, which have planned missing data, have been found to have robust power despite small sample sizes at certain ages (Rhemtulla & Hancock, 2016). Data were log-transformed prior to analysis to better approximate normality. Random intercept and slope factor scores were generated for subsequent biometric analyses using maximum a posteriori (MAP) scoring (MacCallum, 2009). The individual factor scores were obtained from an unconditional, two-level model in which measurement occasions were nested within participants, as biometric twin models account for clustering within families. Biological sex, ethnicity, and ADI were regressed out of the factor scores prior to running the twin analyses (McGue & Bouchard, 1984).

Biometric Twin Analyses

A series of biometric twin models were then run to estimate the relative genetic and environmental influences on variability in the estimated intercept and slope factor scores. We used these factor score estimates rather than running a full biometric latent growth curve model due to the large variation in participant age at each timepoint (time is typically based on assessment schedule in biometric latent growth curve models). Time can be easily modelled via participants' chronological age at a given time point in multilevel growth models, however, making it better-suited for study designs that involve an uneven schedule of assessments (Hox & Stoel, 2014), such as this one. Although each analytic framework has its practical advantages, structural equation and multilevel growth models are conceptually analogous.

Classical twin models leverage the difference in the proportion of segregating genes shared between identical (MZ) twins, who share 100% of their genes, and fraternal (DZ) twins, who share an average of 50% of their segregating genes to estimate the relative contributions of genetic and environmental influences to the variance within observed behaviors (phenotypes). Phenotypic variance is decomposed into three variance components: additive genetic (A), shared environmental (C), and nonshared environmental (E). More information on twin modeling is provided elsewhere (Neale & Cardon, 1992). For the present study, we first computed the A, C, and E estimates for the random intercept and slope factors, respectively. We then used a bivariate twin model to clarify the extent to which the etiologies of the slope and the intercept overlapped.

Results

Descriptive Statistics

Descriptive statistics by age are shown in Table 1.1, while Table 1.2 contains correlations across assessment waves (operationalized here in person-specific chronological order because of the irregularities in when specific assessments were administered). Participants evidenced moderate-to-high rank-order stability in their reported ASB over time, with correlations ranging from .20 to .56. Paired-sample *t* tests further indicated that, within persons, ASB decreased significantly from the first assessment to the second ($t(1913) = -7.97$, $p<.001$), and from the second assessment to the third ($t(1333) = -9.55$, $p<.001$). Changes across subsequent assessments were not significant. In addition, mean ASB scores decreased steadily with age (see Figure 1.1). Collectively, these findings indicate that, while participants who exhibited high levels of ASB during childhood largely continued to do so during adolescence, the absolute level of ASB decreased significantly over time. Although not shown in Table 1.2, males had slightly higher ASB scores than females at the first two timepoints (Cohen's d = .21 and .22, respectively; both $p < .001$), whereas there were no significant sex differences at assessments three, four, or five. Lastly, ADI was significantly correlated with ASB at the first two timepoints (r = .17 and .11, respectively, both $p < .001$), but not at the last three.

Multilevel Modeling

Phenotypic Results

In the baseline unconditional growth model, ASB decreased significantly over time (slope mean = -.03, $p<.001$), although there was significant interindividual variation in the magnitude of the decline (slope variance=.001, $p<.001$). The covariance between the intercept and the slope was negative, meaning that participants with higher ASB scores at baseline tended to exhibit a more rapid decline in ASB over time. We next fitted a conditional growth model. Results are shown in Table 1.3. Sex significantly predicted both intercept and slope, such that male participants had higher ASB scores at baseline and displayed a more rapid decline with age. ADI was also a significant predictor of both the intercept and slope, with participants from more impoverished neighborhoods exhibiting higher levels of ASB at baseline and declining more rapidly over time. By contrast, race (a socially constructed category coded as white/non-white given the composition of our sample) did not significantly predict the intercept or the slope.

Twin Model Results

Univariate ACE models were first run using intercept and slope factor scores from the unconditional growth model, in order to estimate the genetic, shared environmental, and nonshared environmental contributions to stability and change in ASB. Standardized univariate variance estimates are presented in Table 1.4. For both the intercept and the slope, there were significant genetic and nonshared environmental contributions, but no significant shared environmental influences, although the estimated magnitude of the shared environmental variance for the intercept was non-zero (.10).

The bivariate ACE model indicated both significant unique and overlapping genetic and nonshared environmental influences on the intercept and the slope factors. Path estimates are shown in Figure 1.2. More than one-third (38%) of the genetic variance in the slope factor was shared with the intercept. Thus, the genetic etiology of ASB development was due to both genetic influences that had emerged during the preschool years, as well as to novel genetic influences emerging later in development. Interestingly, one-third of the nonshared environmental variance in the slope was also present at baseline, indicating that our estimates of E did not represent solely transient, time-specific influences, but rather exhibited a fair amount of stability across development. Consistent with the univariate results, shared environmental contributions were not significant. Taken together, the genetic and nonshared environmental influences on ASB in early childhood also appear to contribute to its stability across development.

Discussion

The aim of the present study was to elucidate genetic and environmental contributions to continuity and change in ASB from early childhood into emerging adulthood. To do so, we obtained estimates of participants' baseline level of ASB and change over time via multilevel growth curve modeling. We then made use of a series of classical twin models to illuminate genetic, shared environmental, and nonshared environmental contributions to the intercept and the slope of ASB, as estimated via factor scores generated in the prior analyses. The results indicate that initial levels and change over time in ASB were due to both genetic and nonshared environmental influences, some of which overlapped. Neither initial level of ASB nor change over time were subject to significant shared environmental influences.

Genetic influences were found to make important contributions to ASB in early life, as well as to change in ASB across development. Furthermore, more than one-third of the genetic contributions to change over time were already present at baseline (i.e., during the preschool years), indicating a fair amount of continuity in genetic influences throughout early development. These findings are consistent with those of other studies that have found prominent genetic influences on continuity in youth ASB, albeit during much shorter windows of development (Bartels et al., 2004; Burt et al., 2007; Eley et al., 2003; Porsch et al., 2016) or over longer periods that did not include late adolescence/emerging adulthood (Pingault et al., 2015). Our study extends these findings by indicating that genetic influences contribute to continuity in ASB across all of early development (i.e., the first 20 or so years of life). Genetic influences were also found to underlie change in ASB, with nearly two-thirds of the genetic variance in the slope representing novel influences that were not present at baseline. The emergence of novel genetic influences over the course of our study is unsurprising, given the broad age range (3 to 22 years) represented in the sample. In addition, this pattern of results is consistent with those of other studies of children and adolescents that have found genetic influences to contribute to both continuity *and* change in ASB (Bartels et al., 2004), particularly for nonaggressive rule-breaking (Eley et al., 2003).

Nevertheless, neither continuity nor change in ASB were due solely to genetic influences. Nonshared environmental variance played a considerable role in continuity across development in our sample, with fully one-third of the nonshared environmental contributions to the slope already present at baseline. Such findings stand in contrast to those of prior longitudinal studies of youth ASB, which have found the nonshared environment to exert largely transient effects on change over time that were specific to each assessment wave (Bartels et al., 2004; Burt et al., 2007; Eley et al., 2003). Because nonshared environmental influences tend to become more stable with age (Burt et al., 2015; Hopwood et al., 2011), it is possible that our study was better positioned to detect stability in the nonshared environment compared to those conducted in samples of children and young adolescents (e.g., Bartels et al., 2004; Eley et al., 2003). The possibility of nonshared environmental influences contributing increasingly to stability with age is also consistent with developmental theories of canalization, which posit that, as youth begin to shape their own environments, their range of potential outcomes typically narrows. In other words, individuals increasingly follow idiosyncratic trajectories in accordance with both genetic predispositions and environmental exposures that they themselves may seek out (e.g., Turkheimer & Gottesman, 1991). For example, a twin who is parented more harshly during preschool may experience difficulty regulating their emotions throughout childhood and adolescence and increasingly choose to spend time with peers who have similar difficulties, further differentiating the child from their co-twin. That said, this interpretation is not consistent with the findings of Burt et al. (2007), which also identified transient effects of the nonshared environment between late adolescence and early adulthood. However, Burt and colleagues examined diagnostic symptom counts of Antisocial Personality Disorder, a more extreme phenotype than the more dimensional ASB assessment examined here (Lahey et al., 2005).

What might be the specific non-shared environmental experiences that underlie stability in ASB? One possible non-shared environmental influence is delinquent peer affiliation, which has been found to predict growth in ASB throughout adolescence (Eamon, 2002; Gardner et al., 2008). That said, prior twin work has suggested that twin differences in delinquent peer affiliation appear to be a consequence, rather than a cause, of differences in their ASB (Burt, et al., 2009). Another possibility centers on aspects of the family environment that, while objectively shared by siblings, impact each child in idiosyncratic ways (e.g., siblings respond differently to parental divorce) (Goldsmith, 1993). Such familial influences could have an enduring impact on ASB development throughout childhood and adolescence. A final possibility is differential parenting, which may represent a relatively continuous influence that stably differentiates children in the same family. Such considerations are consistent with theoretical work positing that “proximal processes”, or reciprocal interactions between the individual and their immediate environment, play a critical role in shaping behavioral development (Bronfenbrenner, 1988), and empirical work identifying harsh parenting as a risk factor for child, adolescent, and young adult ASB in particular (Beauchaine et al., 2005; Conger et al., 1994; Gard et al., 2017). Furthermore, studies of within-family differences in parental harshness significantly predicted within-pair differences in monozygotic twins’ ASB, both cross-sectionally (Burt et al., 2021) and over time (Burt, et al., 2006). Such findings point to parenting as a particularly promising target for subsequent studies of the environmental etiology of ASB.

Notably, however, only environmental exposures unique to each child in a given family appeared to impact change in ASB across development, as shared or common family-level environmental influences were negligible. While the shared environment has previously been found to contribute to continuity in ASB during childhood and early adolescence (Bartels et al., 2004; Eley et al., 2003), it has not been found to impact ASB development during emerging adulthood (Burt et al., 2007). While our inclusion of emerging adulthood could conceivably contribute to our null findings for shared environmental influences, we also note that ASB was assessed using a 5-item screening measure of youth behavior problems. Brief measures often have lower reliabilities than do longer measures, an especially salient point here since increased measurement error would increase estimates of nonshared environmental effects (see Burt (2009) for a discussion of factors affecting detection of shared environmental effects). Consistent with the latter, supplemental analyses using participants' scores on the full CBCL Conduct Problems scale (17 items) across the TBED-C and MTwiNS administrations indicated that there were significant shared environmental contributions to baseline ASB during middle childhood (C variance estimate = .20, $p < .05$), although shared environmental influences on rate of change remained non-significant (C variance estimate = .11). Nevertheless, there were also significant, and partially overlapping, genetic and nonshared environmental contributions to both intercept and slope for CBCL scores, consistent with our results for the 5-item measure (see Table S1.1 and Figure S1.1 in Supplemental Material).

There are several other limitations to keep in mind when interpreting the results of the present study. First, our analyses are not able to clarify the exact duration of nonshared environmental contributions to ASB. While the significant overlap in these contributions at baseline and over time indicates some degree of continuity, it is unclear whether the influences that do not overlap represent transient effects lasting minutes or days, or more enduring effects that contribute to systematic change. Second, there was a drop in sample size at ages 11-12 and 20-22. As our intercept and slope estimates were based on growth curves, however, there is relatively little impact of ASB estimates at one particular age on participants' overall trajectories.

Next, the SDQ does not delineate aggressive and non-aggressive rule-breaking sub-types of ASB. This is potentially problematic since these two dimensions of ASB have been shown to exhibit distinct etiologies and developmental trajectories (Burt, 2012). Indeed, our finding that ASB decreased linearly across development likely indicates that our measure was unable to capture the spike in rule-breaking typically seen in studies spanning adolescence (e.g., Bongers et al., 2004; Windle, 2000). There is thus a need for subsequent research on the development of rule-breaking and aggression as separate phenotypes from childhood into emerging adulthood, particularly using developmentally sensitive measures that capture differences in symptom presentation by age (i.e., heterotypic continuity). However, the items included on the SDQ, and in our abbreviated scale from the CBCL, screen for behaviors that are typically present, to some degree, throughout early development (e.g., lying, disobedience). Moreover, scores on the 17-item Conduct Problems scale on the CBCL also declined linearly across the three TBED-C/MTwiNS assessments, which spanned ages 6 to 19, indicating that the brevity of our measure likely did not prevent it from capturing age-related trends in ASB development in our sample.

In addition, child sex was entered as a covariate in models including male and female participants. Some longitudinal twin studies (e.g., Burt et al., 2007; Eley et al., 2003) have found models allowing for sex differences in the etiology of ASB development to fit better than models constraining parameters to be equal across sex. However, this pattern of results is generally the exception rather than the rule (Burt et al., 2019; Jacobson et al., 2002). Moreover, we note that Burt et al. (2007) found few differences between male and female participants in standardized parameter estimates. As such, while males evidenced higher levels of ASB at baseline and somewhat more decline over time relative to females in our study, we do not expect our overall conclusions to differ in models allowing for sex differences in parameter estimates.

No differences were observed between white participants and those identifying with marginalized races/ethnicities in either baseline ASB or change over time in our sample. Given the demographics of the State of Michigan, however, there were not sufficient numbers of those who identified with any specific marginalized race or ethnicity to model these groups separately. There was a significant effect of neighborhood disadvantage, with youth from impoverished neighborhoods exhibiting higher levels of ASB at baseline and more rapid decline over time. That less privileged youth had higher initial levels of ASB is consistent with a large body of research demonstrating that familial and neighborhood disadvantage increases risk for nearly all youth psychiatric disorders (e.g., Kupersmidt et al., 1995; Leventhal & Brooks-Gunn, 2000). Our finding that these youth also desisted more quickly suggests that discrepancies in behavioral outcomes by socioeconomic status may decrease with age. Regardless, there is a need for further research examining disadvantage in the broader context (e.g., neighborhoods, schools), and inequitable structural characteristics (e.g., differences in policing, housing policies) in particular, as a predictor of ASB development over time in racially, ethnically, and socioeconomically diverse samples.

Despite these limitations, the present study is the first to examine the genetic and environmental etiology of ASB over time in a sample spanning nearly all of childhood, adolescence, and emerging adulthood. The key strength of such a study, when incorporating a twin design, is its potential to elucidate the genetic and environmental factors contributing to human development across multiple stages of the life course. Our study yielded two important conclusions. First, genetic factors contributed significantly to both continuity and change in ASB. Given the broad age range under study, the genetic contributions to continuity are perhaps more noteworthy. Nearly 40% of genetic influences on change throughout development were already present at the baseline assessment, which was conducted as early as age 3 in some participants. Such findings underscore the importance of genetic influences in shaping ASB trajectories. While the specific genetic factors underlying continuity and change are unknown, one possibility is that genetic contributions to improved behavioral and emotional regulation are activated as youth progress through adolescence, resulting in fewer problem behaviors over time. On the other hand, genetic factors underlying dimensions of temperament that are known to be predictive of ASB, including negative emotionality and disinhibition, may be among those contributing to continuity in ASB, as temperament is both heritable and moderately stable throughout development (Ganiban et al., 2008). Future work should seek to contrast and test these two possibilities.

Second, the nonshared environmental influences on ASB reflected not only transient person-specific environmental influences, but also more enduring influences that overlapped across assessment waves. Put another way, environmental influences unique to each child within a given family, rather than shared exposures affecting the entire family, were found to be important for both stability and change in ASB across early development. Such findings are consistent with research indicating the importance of the nonshared environment to behavioral outcomes (Plomin & Daniels, 1987). Subsequent studies should seek to identify specific nonshared environmental influences that persist over time prior to adulthood.

Tables

Table 1.1. *Descriptive statistics: youth ASB scores by age*

Age	Total *N*	MTP early	TBED-C	MTP late	MTwiNS	Online	Mean ASB (SD); range
3	288	288	0	0	0	0	1.91 (1.63); 0-8
4	284	284	0	0	0	0	1.65 (1.49); 0-8
5	286	286	0	0	0	0	1.74 (1.55); 0-7
6	858	196	600	62	0	0	1.49 (1.62); 0-10
7	584	132	408	38	6	0	1.39 (1.54); 0-10
8	682	254	344	66	18	0	1.37 (1.54); 0-10
9	556	128	338	78	12	0	1.15 (1.42); 0-9
10	476	34	312	92	38	0	1.17 (1.47); 0-7
11	118	0	58	8	36	16	1.01 (1.37); 0-5
12	116	0	0	14	58	44	1.09 (1.43); 0-6
13	200	0	0	4	62	134	.88 (1.31); 0-6
14	290	0	0	2	146	142	1.00 (1.56); 0-10
15	312	0	0	0	150	162	.83 (1.42); 0-9
16	260	0	0	4	96	160	.70 (1.26); 0-9
17	254	0	0	0	78	176	.95 (1.43); 0-9
18	146	0	0	0	18	128	.75 (1.09); 0-5
19	156	0	0	0	2	154	.86 (1.39); 0-9
20	116	0	0	0	0	116	.72 (1.15); 0-5
21-22	44	0	0	2	0	42	.57 (1.07); 0-5

Table 1.2. *Sample sizes and correlations in ASB over time*

	Time 2	Time 3	Time 4	Time 5	Mean age (SD)	Mean ASB (SD)	*N*
	r						
Time 1	**.49***	**.31***	**.30***	.20	6.57 (2.09)	1.56 (1.61)	1996
Time 2	-	**.40***	**.36***	**.46***	8.63 (2.60)	1.28 (1.49)	1978
Time 3		-	**.56***	**.55***	15.29 (3.14)	.87 (1.34)	1354
Time 4			-	**.53***	16.61 (2.16)	.87 (1.41)	551
Time 5				-	16.79 (.86)	.48 (.87)	56

Note: Bold font and asterisk indicate that the correlation was significantly different than zero at p<.05. Time points indicate assessment waves in chronological order, which varies across participants (e.g., Time 1 was MTP 1 for some participants, TBED-C for others).

Table 1.3. *Key parameter estimates from conditional multilevel growth curve model of ASB development*

Parameter	Estimate	S. E.	*p*-value
Fixed effects (means)			
Intercept	**.667**	**.069**	**<.001**
Slope	**-.018**	**.007**	**.010**
ADI → intercept	**.004**	**.001**	**<.001**
Sex → intercept	**.170**	**.032**	**<.001**
Ethnicity → intercept	-.035	.054	.518
ADI → slope	**-.0002**	**(.000)**	**.001**
Sex → slope	**-.008**	**.003**	**.016**
Ethnicity → slope	.001	.006	.900
Random effects (variances)			
Intercept	**.179**	**.015**	**<.001**
Slope	**.001**	**.000**	**<.001**
Intercept-slope covariance	**-.010**	**.001**	**<.001**
Residual variance	**.152**	**.006**	**<.001**

Note: Bold font indicates that the estimate was significantly different than zero at $p<.05$.

Questionnaire type was included as a random, time-varying covariate.

Table 1.4. *Standardized variance estimates from univariate ACE model*

	Variance estimates		
	a	c	e
Univariate			
Intercept	**.53***	.10	**.37***
Slope	**.45***	.03	**.52***

Note: Bold font and asterisk indicate that the estimate was significantly different than zero at p<.05.

Figures

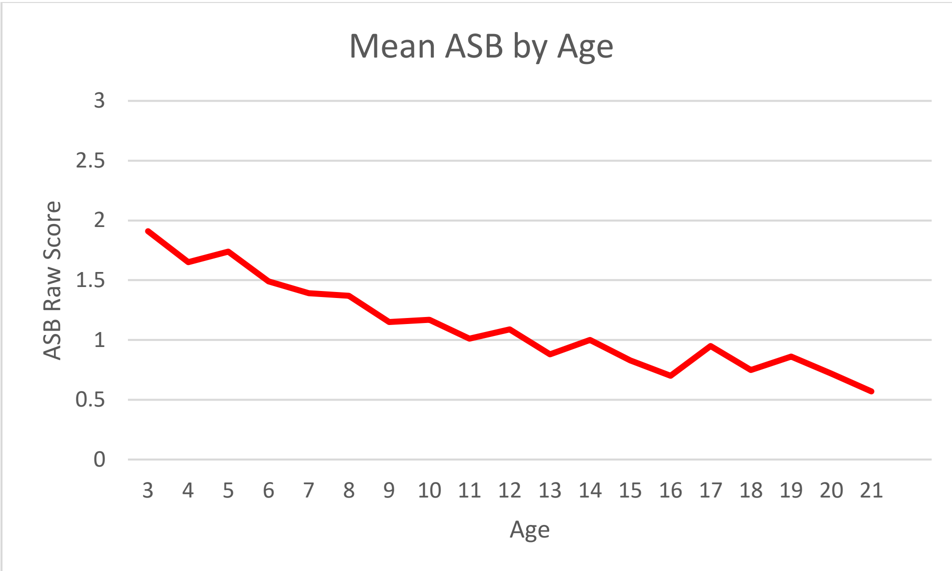


Figure 1.1. *Age-related change in ASB.*

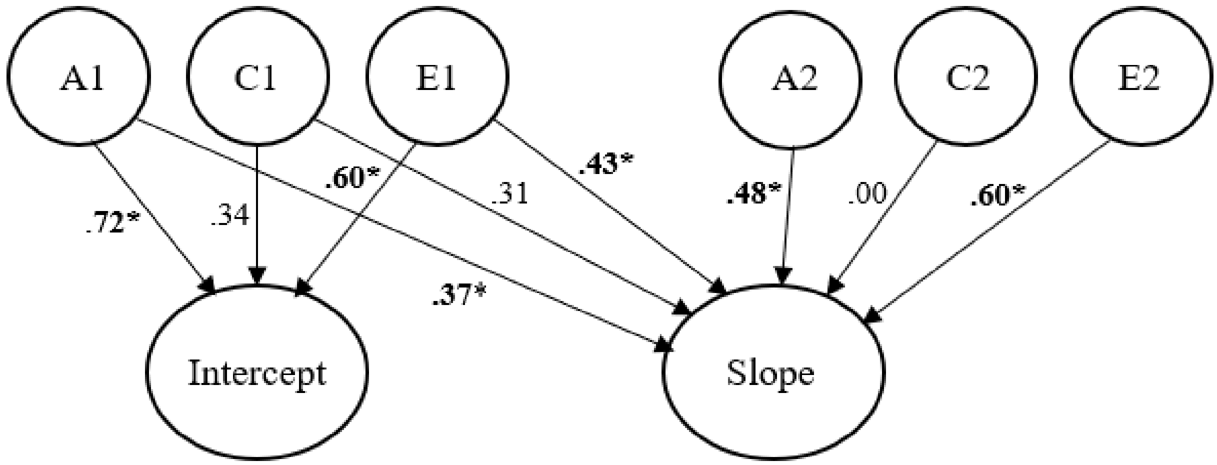


Figure 1.2. *Path diagram of a bivariate twin model.* The variance in the intercept and the slope is partitioned into additive genetic effects (A1 and A2), shared environmental effects (C1 and C2), and nonshared environmental effects (E1 and E2). For ease of presentation, this path diagram represents one twin in a pair. Standardized path estimates are squared to represent the proportion of variance accounted for.

CHAPTER 2: UNDERSTANDING DESISTANCE FROM AGGRESSION: AN EMPIRICAL INTEGRATION OF PERSON-CENTERED AND VARIABLE-CENTERED APPROACHES

Abstract

When leveraged together, variable-centered and person-centered statistical methods have the potential to illuminate the factors predicting mental health recovery. However, because extant studies have largely relied on only one of these methods, we do not yet understand why some youth demonstrate recovery while others experience chronic symptoms. This omission limits our understanding of trajectories of physical aggression (AGG) in particular, which are frequently characterized by desistance. The present study examined the development of AGG across childhood and adolescence via variable-centered and person-centered modeling, with neighborhood and family characteristics considered as moderators. Variable-centered results indicated a mean-level decline in AGG with age but were less useful for illuminating predictors of that decline. Person-centered analyses, by contrast, identified low parent-child conflict and high household income as predictors of desistance. Although variable-centered analyses were integral to modeling the average AGG trajectory, person-centered techniques proved more useful for understanding predictors of desistance.

Introduction

Physical aggression (AGG) encompasses a broad spectrum of behaviors that violate the personal rights of others, ranging from relatively minor acts (e.g., hitting, kicking) to more serious violent crimes (e.g., stabbing, shooting) that bring offenders into contact with the criminal justice system. Trajectories of AGG are characterized by both high rank-order stability over time and substantial changes with age. A large body of research, for example, has found that the frequency of AGG peaks during the preschool years followed by steady declines across childhood and adolescence (e.g., Murray et al., 2020; NICHD Early Child Care Research Network & Arsenio, 2004), such that desistance increasingly becomes the norm (e.g., Broidy et al., 2003). For those youth who do not desist, however, their trajectories often culminate in a series of poor adult outcomes (e.g., poor mental and physical health, financial difficulties, incarceration; Odgers et al., 2008).

There is thus a clear need to illuminate how AGG develops over time and to identify predictors that interrupt chronic AGG trajectories. A substantial body of research has sought to do the former using either variable-centered or person-centered approaches to longitudinal data analysis. Variable-centered studies have generally indicated that mean-level increases are normative during the preschool years, but that desistance becomes the norm as children progress through school. Longitudinal person-centered studies, in turn, have typically identified three or four trajectories, with the modal group following a declining trajectory by middle childhood (Carroll, Mikhail, & Burt, 2023). Each statistical approach also provides distinct, yet complementary, information about the predictors of AGG. For example, most variable-centered studies have found males to exhibit greater AGG than females at baseline but comparable rates of change, whereas person-centered studies have indicated that males are typically over-represented in the high-risk trajectories. Likewise, hostile/conflictual family interactions, familial socioeconomic disadvantage, and neighborhood deprivation have each been found to predict greater baseline levels of AGG in variable-centered models (e.g., Benson & Buehler, 2012;

Olson et al., 2013; Sacco et al., 2015), as well as membership in “chronic” person-centered trajectory groups (e.g., Côté et al., 2006; Spano, Rivera, & Bolland, 2010).

However, all of the above longitudinal studies relied on only one of the two statistical approaches (variable-centered or person-centered), with key downstream consequences for our understanding of desistance. Variable-centered approaches use growth curve modeling to quantify the mean rate of change over time for the entire sample, which is thought to comprise a single population (see Figure 2.1a) (McArdle & Epstein, 1987). Individual patterns of continuity and change are captured by the intercept, or level of the outcome variable at a particular point in time (often at baseline), and the slope, or rate of change. The means of the intercept and slope represent the average pattern of development across the full sample, whereas the variances represent individual differences in the initial level of the outcome and rate of change, respectively (Preacher, Wichman, MacCallum, & Briggs, 2008). For example, a variable-centered analysis of AGG might find the mean linear slope to be negative and the intercept and slope variances to each be significant. These findings would indicate that the average participant exhibited a steady decline in AGG with age, but that there were significant between-person differences in both initial level of AGG and rate of change.

Person-centered approaches, by contrast, consider the sample under study to comprise distinct subgroups that differ by intercept and/or slope. Many person-centered studies of AGG have used a technique developed by Nagin (1999), termed semi-parametric mixture modeling or latent class growth analysis (LCGA). In this approach, participants are assigned to latent, or unobserved, trajectory groups based on their posterior probability of group membership (i.e., their likelihood of belonging to a particular group given their scores on the variable of interest) (Figure 2.1b). For example, a participant who consistently reports high levels of AGG would have a high probability of assignment to a “chronic” trajectory group, whereas a participant who reports frequent engagement only at baseline would likely be assigned to a “desisting” trajectory. Although groups derived via LCGA are free to vary from one another by intercept and slope, in most studies, participants within the same group are assumed to follow the same trajectory (i.e., intercept and slope variances are fixed within groups).

Variable-centered and person-centered approaches are each able to provide only a *partial* picture of the development of AGG (Carroll, Mikhail, et al., 2023). The growth curve representing average change in AGG in variable centered studies, for example, may not adequately capture the developmental heterogeneity observed in key participants. That is, even if the mean slope indicates a modest decline over time, some youth might increase in their delinquency over time while others may fully desist. Put differently, there may be more information in the variance than in the mean slope itself. The primary strength of person-centered methods lies in their ability to illuminate the slope variance, identifying specific groups of youth whose patterns of engagement make them important targets for intervention efforts. However, person-centered approaches are critiqued for overextraction, or identifying latent classes that do not actually exist in the sample (Bauer & Curran, 2004). Indeed, some researchers argue that, because AGG varies continuously across the population, it has no qualitatively distinct underlying classes and is best-represented using variable-centered models (e.g., Walters & Ruscio, 2013). In turn, proponents of person-centered methods contend that the groups serve as an approximation of participants' trajectories and need not exist in reality to be theoretically and empirically useful (Nagin & Tremblay, 2005). Each technique thus has clear advantages and important drawbacks.

In short, although both person-centered and variable-centered approaches provide important information, their respective conclusions are necessarily limited by the particulars of their analytic approach. One obvious solution to this dilemma is to conduct both sets of analyses in the same dataset with the goal of jointly interpreting their findings. What might we gain from doing this? Variable-centered studies illuminate normative patterns of development across entire samples and are positioned to identify factors distinguishing between initial levels of the trait under study, as well as factors predicting rate of change across the entire sample.

However, they are less helpful for examining change among those youth who demonstrate high levels of aggression in childhood and subsequently desist. Person-centered techniques are uniquely positioned to identify these youth, whose trajectories are of particular interest to most clinical psychologists. Moreover, person-centered analyses can not only distinguish between youth exhibiting chronic behavior problems and their counterparts who desist, but also identify risk and protective factors that predict membership in the desisting versus persisting classes. Such methods may also facilitate precision medicine by allowing interventions to be tailored to youth following similar trajectories.

A recent review (Carroll, Mikhail, et al., 2023) began the process of integrating variable-centered and person-centered approaches through a conceptual review of the developmental literatures on antisocial behavior. Of the 124 studies of AGG included in the review, only four (3%) examined variable and person-centered techniques simultaneously (Dong, 2016; Ehrenreich, Beron, Brinkley, & Underwood, 2014; Ingoldsby, 2003; NICHD Early Child Care Research Network & Arsenio, 2004). The results of these studies yielded little information about the predictors of desistance from AGG. Two did not examine predictors of the slope variance (Dong, 2016; Ehrenreich et al., 2014), and one was under-powered (Ingoldsby, 2003; N = 170). Only one study identified specific predictors of desistance, which included higher maternal educational attainment and fewer maternal depressive symptoms, but participants were not followed past age nine (NICHD & Arsenio, 2004). Moreover, no study to date has employed both person-centered and variable-centered techniques to examine predictors of desistance from AGG across childhood and adolescence. Such studies would be crucial for identifying targets to leverage in intervention efforts for youth demonstrating elevated levels of AGG, who comprise a sizeable proportion (~10% to >50%; e.g., Broidy et al., 2003; Nagin, Barker, Lacourse, & Tremblay, 2009) of childhood samples.

Present Study

The present study sought to illuminate the development of AGG from middle childhood through late adolescence in a large, population-based sample enriched for neighborhood disadvantage. The development of AGG was modeled via a series of variable-centered and person-centered analyses, and biological sex, household income, parenting, and neighborhood disadvantage were examined as moderators of AGG growth curves and trajectories. We examined socioeconomic disadvantage at both the family and neighborhood levels, in light of prior research indicating that each form of disadvantage has unique implications for youth behavioral outcomes, particularly externalizing (Carroll et al., 2023). Consistent with prior work (e.g., Brame, Nagin, & Tremblay, 2001; Lee, Liu, & Watson, 2016), we hypothesized that 1) the mean level of AGG would decline across the study period, 2) there would be either three or four AGG trajectory groups, and 3) female sex, higher household income, greater parental nurturance, lower parent-child conflict, and greater neighborhood advantage would predict lower levels of AGG at baseline as well as desistance from AGG across adolescence.

Transparency and Openness

The present study was not preregistered. Because of the language in the informed consent document at intake, we cannot post the data publicly, but they can be obtained from the primary author upon reasonable request. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All procedures were approved by the primary author’s institutional review board and are in compliance with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration. Child participants provided informed assent, and parents provided informed consent for themselves and their children.

Methods

Participants

The population-based Michigan State University Twin Registry (MSUTR) includes several independent twin projects. Participants in the current study were drawn from the Twin Study of Behavioral and Emotional Development in Children (TBED-C), a study within the

MSUTR. The TBED-C includes a population-based arm (*N*=528 families), and an ‘under-resourced’ arm for which inclusion criteria also specified that participating twin families lived in neighborhoods with neighborhood poverty levels at or above the Census mean at study onset (10.5%) (*N*=502 families). To recruit families at intake, the Department of Vital Records in the Michigan Department of Health and Human Services (MDHHS) identified twins in our age-range via the Michigan Twins Project (MTP), a population-based registry of more than 32,000 twins in Lower Michigan recruited via birth records. The Michigan Bureau of Integration, Information, and Planning Services database was used to locate family addresses no more than 120 miles from East Lansing, Michigan through parent drivers’ license information. Pre-made recruitment packets were then mailed to parents by the MDHHS. Parents who did not respond to the first mailing were sent additional mailings roughly one month apart until either a reply was received or up to four letters had been mailed.

This recruitment strategy for the TBED-C yielded an overall response rate of 57% for the at-risk sample and 63% for the population-based sample. Other recruitment and sampling details can be found in prior publications (i.e., Burt & Klump, 2019). The two arms of the study were analyzed jointly for the current analyses and were 48.7% female and 51.3% male. Across the full sample, participants endorsed racial/ethnic identities in the following proportions: white (non-Latinx), 82%; Black, 10%; Latinx, 1%; Asian, 1%; Indigenous, 1%; multiracial, 6%. However, families in the under-resourced arm, but not the population-based arm, were more racially diverse than the local population (e.g., 14% Black and 77% white in the under-resourced arm versus 5% Black and 87% white in the population-based arm; 5% Black and 85% white in the local area Census).

Because 90+% of TBED-C families were recruited out of the MTP, we were able to use the MTP data to compare families who chose to participate in TBED-C with those who were recruited but did not participate. TBED-C families were generally representative of recruited but non-participating families. As compared to non-participating twins, participating twins reported similar levels of conduct problems, emotional symptoms, and hyperactivity (*d* ranged from -.08 to .01 in the population-based arm and .01 to .09 in the under-resourced arm; all ns). Participating families also did not differ from non-participating families in paternal felony convictions (*d* = -.01 and .13 for the population-based and under-resourced arms, respectively), rate of single parent homes (*d* = .10 and -.01 for the population-based and under-resourced arms, respectively), paternal years of education (both $d \leq .12$), or maternal and paternal alcohol problems (*d* ranged from .03 to .05 across the two arms). However, participating mothers in both samples reported slightly more years of education (*d* = .17 and .26, both $p < .05$) than non-participating mothers. Maternal felony convictions differed across participating and non-participating families in the population-based arm (*d* = -.20; $p < .05$) but not in the under-resourced arm (*d* = .02). In short, our recruitment procedures appear to have yielded a sample that was representative of both recruited families and the general population of the State of Michigan.

Behavioral data regarding the children's AGG were collected at up to three time points. All 1,030 families were assessed once in middle childhood (ages 6-11; M_{age} (SD)=8.02 (1.49)) as part of the TBED-C (Wave 1). TBED-C participants residing in modestly-to-severely disadvantaged neighborhoods (*N* = 768 families) are currently being recruited for reassessment as adolescents up to two times, 18-months apart, through the Michigan Twin Neurogenetics Study (MTwiNS). For the present study, the sample sizes at Waves 2 and 3 were 760 and 414, respectively.

Procedure

TBED-C families completed their intake assessments between 2008 and 2015. Assessment teams consisted of two research assistants and at least one paid staff member and took 4-5 hours to complete (lunch was provided). Families completed questionnaires, interviews, and videotaped interactions during the in-person assessment. The primary caregiver (nearly always the mother) participated alongside the twins. Assessments typically took place in our university laboratory (903 families). If families were unable or unwilling to travel, however, assessments took place in participants' homes (13%). Families with older twins were less likely to complete home visits than on-site assessment visits (Cohen's d=-.24, p=.01), as were families that identified as White (d=-.21; p<.05). Similarly, families with fewer financial resources were more likely to participate in home visits versus on-site assessments (d=-.32; p=.004). However, families completing on-site assessments versus home visits did not differ in maternal education, zygosity of the twins, or sex of the twins (*d*s ranged from -.02 to .16, all ns). The MTwiNS follow-up assessments were also completed in-person. The twins and their primary caregiver completed an in-person assessment lasting 4-8 hours at either the East Lansing or Ann Arbor-based laboratories.

Measures

AGG

Youth AGG was defined via parent report. The primary caregiver (nearly always the twins' mother) completed the 18-item AGG scale from the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) at all three waves, rating the extent to which a series of statements described the child's behavior over the past six months using a three-point scale (0=never to 2=often/mostly true). Caregiver-informant reports were available for 99% of participating twins at Wave 1, 99% at Wave 2, and 90% at Wave 3. One issue with the AGG scale from the Achenbach family of instruments is that the items assess not only physically aggressive behaviors (e.g., fights, destroys belongings) but also verbal AGG (e.g., argues) and emotional dysregulation (e.g., sudden changes in mood). Because the aim of our study was to model the development of physical AGG, we constructed a shortened scale of items from the Achenbach instruments via a series of item response theory (IRT) analyses (see Supplementary Methods in Appendix B).

Familial Context

Familial socioeconomic deprivation was assessed via maternal reports of annual household income at the TBED-C assessment. Household income was measured on a 10-point Likert scale (1 = <$10,000 to 10 = >$50,000).

At the TBED-C assessment, participating mothers and children also completed the Parental Environment Questionnaire (PEQ; Elkins et al., 1997), which assesses several dimensions of the parent-child relationship using a 4-point Likert scale ranging from "definitely false" to "definitely true". We focused on the Conflict and Nurturance scales (12 items each; $\alpha \geq .68$ for all scales and informants). The Conflict scale assesses harsh or conflictive parenting practices (e.g., "I often criticize my child"), while the Nurturance or Involvement scale assesses parental communication, support, and involvement with their child (e.g., "I praise my child when he/she does something well"). Items are nearly identical across parent and child versions of the PEQ, with minor alterations in wording. Responses were coded so that higher scores indicated higher levels of each construct. Scores were averaged across informants to create composite reports of conflict and nurturance, respectively. Prior work has found the Conflict and Nurturance scales from the PEQ to demonstrate strong internal consistency reliability for parents and children and to be highly correlated with scores on the Family Environment Scale, which also assesses quality of the parent-child relationship (Elkins et al., 1997).

Neighborhood Context

Neighborhood disadvantage was defined via the Area Deprivation Index (ADI; Kind & Buckingham, 2018), which comprises 17 indices of Census block group disadvantage (e.g., poverty rate, income disparity). Data were weighted according to the factor loadings identified by Kind & Buckingham (2018), and weighted variables were summed to create a deprivation index score for each Census block group. Families were assigned a percentile indicating the level of deprivation in their block group relative to that of all U.S. block groups. Mean ADI scores were 57.24 (SD=22.67; range: 2-99), 61.57 (SD=20.94; range: 6-99), and 60.71 (SD=21.62; range: 12-99) at the TBED-C and MTwiNS assessments, respectively. These scores correspond to a moderate level of neighborhood disadvantage, albeit with considerable variation across participating families. At the TBED-C assessment, for instance, Census tract poverty rates ranged from 0 to 93%, with a mean of 19%.

Race/Ethnicity

Given inequalities in exposure to socioeconomic disadvantage across racial/ethnic identities (Peterson & Krivo, 2009), race was included as a covariate in all longitudinal analyses, coded dichotomously as white/racially marginalized due to the composition of our sample.

Data Analytic Strategy

Sum scores for parent-reported AGG were computed based on the items retained in the final IRT model (see Supplementary Methods in Appendix B for additional details). These scores were log-transformed to adjust for positive skew and were subsequently examined as the outcome measures in the variable-centered and person-centered analyses. Continuous covariates (i.e., income, parenting, and neighborhood disadvantage) were standardized for ease of interpretation.

Variable-Centered Analyses

Analyses were conducted in M*plus* 8.4 (Muthén & Muthén, 1998-2019) using a robust maximum likelihood estimator and full information maximum likelihood estimation (FIML) to account for missing data on the outcome. We examined phenotypic changes in the frequency of AGG over time across the full sample via latent growth curve modeling, which identifies within-person changes in the outcome of interest as well as between-person variation in these changes (McArdle & Epstein, 1987). First, unconditional models were run to determine whether engagement in AGG increased, decreased, or remained constant with age. Age was examined in lieu of assessment wave as the metric of time given the wide range of ages included at each assessment (see Table S2.1). The definition variable approach was used to scale time in years, with the factor loadings constrained to vary by age (Sterba, 2014). Because the definition variables must be complete for the model to be estimated correctly, placeholder values were imputed to represent age for participants with missing data (Grimm, Ram, & Estabrook, 2017).

Specifically, the mean number of years that elapsed between waves was added to participants' ages at Wave 1 if data were missing at Waves 2 and/or 3. Age was available for all participants at Wave 1. Imputed values had no impact on estimation, as missing data on the definition variables was connected with missing data on the outcomes.

Intercepts were centered to represent AGG at age six, the youngest age at the first TBED-C assessment wave. The slope factor represented annual change. Because participants were assessed on only three occasions, we did not examine non-linear growth. Initial analyses compared the respective fits of an intercept-only ("no-growth") model and a model allowing for linear growth. The former had three parameters (intercept mean, intercept variance, and residual variance), and the latter had six (intercept and slope means, intercept and slope variances and their covariance, and residual variance). Model fit was evaluated with four likelihood-based indices: the chi-square difference test (Bollen, 1989), Akaike information criterion (AIC; Akaike, 1987), Bayesian information criterion (BIC; Raftery, 1995), and sample-size adjusted Bayesian information criterion (SABIC; Sclove, 1987). For all indices, lower values indicated better model fit. In the unconditional models, the CLUSTER command was used to account for the nesting of twins within families.

We selected the best-fitting model for the conditional growth curve analyses. AGG was fit conditional upon sex, race, income, conflict, nurturance, and ADI scores, all of which were examined as time-invariant predictors. Because M*plus* employs listwise deletion for missing covariates, all available data were analyzed in the conditional growth curve model by using multiple imputation (Little & Rubin, 2019), with ten imputed datasets. As a result, estimates of the parameters and standard errors reflect the uncertainty that is due to missing data. In M*plus*, it is not possible to include the CLUSTER command while using multiple imputation. Thus, in the conditional LGM, the data were structured to be in "family-wide" format, with parameters constrained to equality across both twins in a pair. The conditional LGM also accounted for the difference in within-pair resemblance for monozygotic (identical) and dizygotic (fraternal) twins, as identical twins are expected to resemble one another more so than fraternal twins on virtually every trait (Turkheimer, 2000).

Person-Centered Analyses

Latent class growth analyses were conducted in M*plus* 8.4 using a robust maximum likelihood estimator and FIML to account for missing outcome data. We examined participants' growth in AGG as a function of their age at each assessment wave. We began with a one-class model and progressed through to a six-class model, consistent with the steps outlined by Muthén (2004) and with the approach employed in prior studies examining trajectories of antisocial behavior (e.g., Odgers et al., 2008). All models allowed the groups to differ from one another by intercept and slope but constrained within-group intercept and slope variances to zero. The optimal model was determined via consideration of two fit indices, the AIC and BIC, as well as the interpretability of the class solution (i.e., the percentage of the sample assigned to each class). Biological sex, race, income, parenting, and ADI were subsequently examined as time-invariant predictors of class membership, using ten imputed datasets to account for missing values. Notably, it is not possible to account for clustering when examining age as the metric of time in M*plus* mixture models, nor was it feasible to examine the data in "family-wide" format because both twins in a pair would be assigned to the same trajectory by default. Thus, we did not account for clustering in our main analyses but ran supplemental analyses with one twin randomly selected per family, discussed in greater detail later in the manuscript.

Results

Descriptives

Descriptive statistics and correlations for the AGG sum scores are shown in Table 2.1. Participants evidenced moderate rank-order stability in AGG over time. Moreover, paired-sample *t* tests indicated that AGG decreased in frequency from Wave 1 to Wave 2 ($t(747) = 13.78$, $p < .001$), and from Wave 2 to Wave 3 ($t(366) = 2.13$, $p < .05$). At the first wave, males had higher levels of AGG than females (Cohen's $d = .285$; $p < .001$); although in the expected direction, sex differences were not significant at Waves 2 or 3. In addition, ADI, income, and nurturance were significantly correlated with AGG at the first two time points, but not at the third. Conflict was significantly correlated with AGG at all waves. All correlations were in the expected direction. (See Table S2.1 for descriptives by age, rather than wave).

Variable-Centered

Unconditional models

Initial growth curve analyses tested the hypothesis that the frequency of AGG decreased with age by comparing the respective fits of an intercept-only model and a linear growth model. The linear growth model fit better than the intercept-only model according to all fit indices (see Table 2.2) and was thus selected as the final unconditional model. At age six, the average participant had a score of .810 on AGG (SE = .022, $p < .001$). On average, scores decreased by .051 points per year (SE = .003, $p < .001$), consistent with hypotheses. There was significant between-person variation in baseline level ($b = .379$, SE = .030, $p < .001$) and in rate of change ($b = .003$, SE = .001, $p < .001$). The estimated covariance between the intercept and slope was -.026 (SE = .004, $p < .001$), indicating that participants with higher AGG scores at baseline declined more rapidly over time.

Conditional model

Sex, race, income, conflict, nurturance, and ADI were subsequently examined as predictors of continuity and change in AGG. As hypothesized, male sex, low income, parent-child conflict, and ADI predicted greater AGG at baseline (see Table 2.2). Race and nurturance predicted neither baseline AGG nor rate of change. The only significant predictors of the slope were sex and parent-child conflict, with males and those with greater conflict demonstrating a more rapid age-related decline.

Person-Centered

Unconditional models

Initial analyses compared the respective fits of a one-class model through to a six-class model to test the hypothesis that there would be either three or four distinct AGG trajectories. Model fit improved as the number of classes increased, as shown in Table 2.3. However, the four through six-class solutions each contained groups that did not appear to follow meaningfully distinct trajectories. For example, the four-class solution comprised two stable trajectories, one of which had a somewhat higher baseline level than the other, whereas the five and six-class solutions each contained two groups that followed desisting trajectories. The three-class solution was thus selected as the optimal model based on fit and clinical relevance. As shown in Figure 2.2, the three-class model comprised a low, stable trajectory representing the majority (53%) of the sample, as well as two trajectories characterized by high levels of AGG at baseline. One of the latter trajectories (33.2% of the sample) followed a desisting pattern with age, whereas the other group (14.1%) continued to aggress at high levels.

Conditional model

Male sex, high ADI, low income, and high conflict predicted membership in the persisting group relative to the low, stable group. Likewise, income and conflict distinguished between desisting and persisting youth. Namely, youth with high levels of AGG in childhood who resided in more affluent households, as well as those exposed to less conflict, were more likely to desist from AGG over the course of adolescence (see Table 2.3 for conditional model results and Table 2.4 for trajectory group characteristics).

Interpretive Integration

Variable-centered and person-centered analyses both indicated a decline in AGG with age. Moreover, male sex, parent-child conflict, low household income, and neighborhood disadvantage all predicted elevated AGG at baseline, evidenced by their prediction of the intercept (variable-centered) and their distinction between the normative group and the elevated, persistent group (person-centered). However, variable-centered and person-centered results differed with respect to predictors of the rate of change. In the former, only sex and mother-reported conflict predicted the slope, and results were not in the expected direction (i.e., male sex and greater conflict predicted *faster* decline with age). Such findings may reflect a statistical artifact, as males and those with greater conflict evidenced higher baseline levels of AGG and thus had more room to decline over time. In the person-centered analyses, by contrast, lower levels of conflict, as well as familial affluence, predicted desistance from AGG among youth demonstrating elevated engagement at baseline. As the desisting and persisting trajectory groups differed primarily by slope, such findings implicate parent-child conflict and familial disadvantage as predictors of the slope variance. Moreover, they point to limitations of variable-centered approaches in understanding heterogeneity in rates of change.

Discussion

The aims of our study were to elucidate the predictors of desistance from AGG across middle childhood and adolescence by integrating findings across two distinct statistical modeling approaches. Consistent with our hypotheses, the modal pattern was one of declining AGG with age. That is, variable-centered analyses revealed a mean-level decline in AGG across the study period, and all three trajectories identified via person-centered modeling followed either stable or decreasing patterns. These findings are consistent with a large body of literature indicating that engagement in AGG typically begins to decline by school entry and continues to do so throughout adolescence (e.g., Broidy et al., 2003; Ehrenreich et al., 2014), facilitated by age-related improvements in emotion regulation and behavioral inhibition (Burt, 2012). The number and shape of the trajectory groups we identified were also largely consistent with findings from studies spanning middle childhood and adolescence, with a modal trajectory characterized by little AGG and two higher-risk trajectories with differing degrees of desistance (e.g., Becht et al., 2016; Maughan et al., 2000).

Although findings were largely consistent across variable and person-centered analyses, each approach provided unique information about the development of AGG, particularly in relation to associations with the predictors. The variable-centered analyses revealed that baseline levels of AGG differed as a function of sex, mother-child conflict, familial socioeconomic status, and neighborhood deprivation. However, only sex and mother-child conflict were observed to predict the slope, and with small effect sizes. Based on the variable-centered findings, one might conclude that youth largely followed similar patterns of declining AGG with age, albeit with some variation between-persons in rate of decline. The person-centered analyses, however, painted a different picture. The trajectories differed not only by baseline engagement but also in their developmental patterns, with one group accounting for most of the age-related decline observed across the full sample.

Perhaps most noteworthy among the person-centered findings is that a sizeable proportion of participants with elevated AGG at baseline had desisted by early adulthood. This finding raises a question that person-centered approaches are uniquely positioned to answer, namely, *which specific factors interrupt trajectories of elevated AGG*? In contrast to the variable-centered findings, the factors predicting patterns of change in the mixture models did so in a way that was consistent with theory. Specifically, familial socioeconomic advantage, as well as parent-child relationships characterized by little conflict, predicted desistance among youth demonstrating elevated AGG at baseline. Not only do these findings suggest potential targets for intervention, but they also point to the utility of person-centered approaches to the study of AGG (and other forms of psychopathology) even within dimensional data. Because variable-centered analyses examine development across entire samples, they are not positioned to extract subgroups that differ widely in their patterns of change, or to identify specific environmental factors predicting such different patterns. Similarly, although person-centered models of antisocial behavior subdivide a continuous trait into distinct categories (e.g., Walters & Ruscio, 2013), the present findings indicate that such approaches can yield information about developmental patterns in continuously distributed data that variable-centered techniques cannot.

We next discuss each of the salient predictors in more detail. As hypothesized, socioeconomic advantage predicted less engagement in AGG. In the growth curve model, youth from wealthier households and non-deprived Census block groups demonstrated lower levels of AGG at baseline than their peers from impoverished backgrounds. Likewise, in the person-centered analyses, high income and low ADI both predicted membership in the stable/low AGG trajectory relative to the persistent one. These findings are consistent with theoretical work pointing to numerous environmental contexts (e.g., family, school, neighborhood) that impact child development (Bronfenbrenner, 1988). They also extend prior, cross-sectional research implicating socioeconomic disadvantage as a multi-faceted construct, with non-interchangeable effects from the family environment and the broader neighborhood context (Carroll et al., 2023; Kupersmidt et al., 1995). Youth who reside in homes and neighborhoods equipped with adequate resources are thus more likely to exhibit low levels of AGG. Moreover, even if they do develop AGG during childhood, those from affluent homes are far more likely to desist than their peers from impoverished backgrounds.

Aspects of the parent-child relationship also predicted trajectories of AGG. Specifically, low levels of mother-child conflict predicted low AGG at baseline in the variable-centered analyses, consistent with prior studies of parent-child conflict and youth behavior problems (El-Sheikh & Elmore-Staton, 2004; Xu et al., 2021). Likewise, in the person-centered analyses, low levels of conflict predicted both low and desisting AGG trajectories relative to persistent AGG. Notably, conflict was one of only two variables that predicted the slope of the latent growth curve, and results were inconsistent with those of the person-centered analyses, such that the former found higher levels of conflict to predict a *faster* decline and the latter found lower levels of conflict to predict desistance. The unexpected findings in the variable-centered analyses may reflect regression to the mean among participants demonstrating high levels of AGG at baseline. That is, participants reporting higher levels of conflict also demonstrated greater baseline AGG and thus had more room to decline with age. The prospect of certain variable-centered findings reflecting regression to the mean artifacts is consistent with prior simulation results. Marsh & Hau (2002), for example, simulated test score data in which schools differed by average achievement at baseline but followed identical growth trajectories. Nevertheless, estimates from multilevel growth curve analyses indicated that schools with lower initial averages demonstrated greater growth in scores over time. An analogous pattern was observed in the present data, namely that participants with higher initial AGG scores, including those with elevated parent-child conflict, appeared to decline more with age.

By contrast, when we examined predictors of trajectory group membership, results for conflict were in the expected direction, with lower conflict predicting both consistently low and desisting trajectories. Because participants in the desisting group had greater conflict than those in the normative group, who had little room to decline with age, higher conflict was associated with more rapidly declining AGG in analyses that collapsed across the full sample. Although reflective of the high levels of conflict among participants with elevated AGG at baseline, variable-centered findings thus provided a somewhat misleading picture of the factors predicting the slope variance. Integration of the variable-centered and person-centered results demonstrates that greater conflict predicts elevated AGG at baseline *and* continued engagement in AGG across adolescence.

Contrary to our hypotheses, however, nurturance did not serve as a protective factor. Prior work has found parental nurturance to have a small association with physical AGG, but at different ages for boys and girls (Arim, Dahinten, Marshall, & Shapka, 2009) or only in conjunction with early pubertal timing in girls (Mrug et al., 2008). Overall, the present findings suggest that negative aspects of the parent-child relationship may be more closely tied to the development of AGG than are positive aspects of the parent-child relationship.

As hypothesized, males demonstrated higher levels of AGG at baseline and were more often assigned to the persistent trajectory group relative to the normative group. Sex also predicted the slope of the growth curve, with males exhibiting a somewhat faster decline in AGG with age, which may again reflect regression to the mean (although males did not exhibit lower levels than females on average prior to age 18; see Figure S2.1). Notably, however, sex did not distinguish between desisting and persisting trajectory group membership, as both groups were predominantly male. Such findings are fully consistent with the broader literature, which has found AGG to be highly sex-specific throughout early development, with some studies identifying up to 15-fold higher rates of AGG for males relative to females (e.g., Berkout, Young, & Gross, 2011).

Limitations

Our results should be interpreted in light of several limitations. First, participants were predominantly white. Although our sample is socioeconomically diverse and representative of the racial/ethnic makeup of the State of Michigan, future studies should examine whether our findings generalize to youth from more diverse racial/ethnic backgrounds as well. Next, participants were assessed on up to three occasions, meaning it was not possible to model non-linear changes in AGG. This may be particularly salient when modeling the development of AGG for adolescents in under-resourced contexts, whose engagement often fluctuates more than that of their wealthier peers (e.g., Ehrenreich et al., 2014; Karriker-Jaffe, Foshee, Ennett, & Suchindran, 2008). Our study was not positioned to detect such fluctuations, although our finding that neighborhood and familial deprivation both predicted persistent AGG is consistent with prior findings that youth exposed to disadvantage are less likely to follow a steadily declining AGG trajectory.

Next, our measure assessed the frequency, not the severity, of AGG, and thus did not detect the age-related increases in severity reported in prior research (e.g., Copeland, Miller-Johnson, Keeler, Angold, & Costello, 2007). In addition, items assessed behaviors of relatively low severity (e.g., hot temper, mean) rather than violent crime, which may explain why we did not identify a “late-onset” or “adolescent-increasing” group in the person-centered analyses. Studies identifying groups with onset after childhood typically examined serious acts of violence (e.g., shooting, stabbing) as the outcome and did not assess for the presence of less serious forms of AGG (e.g., Mata & van Dulmen, 2012; Tung & Lee, 2017). In short, our finding that

AGG was most frequent at study onset (i.e., age 6) is consistent with the broader AGG literature, which has found that AGG rarely first emerges during adolescence or adulthood (Carroll, Mikhail, et al., 2023). Future studies are needed to elucidate associations between familial and neighborhood risk factors and the specific types of AGG behaviors comprising very high-risk trajectories.

Next, we focused on maternal reports of AGG, as they were available at all waves. Although such an approach eliminates confounding due to differences between participants in the number of available informants (i.e., teacher-report data were available for most, but not all, participants at the first two waves), there are inherent limitations to focusing on a single informant. For example, parents may not be aware of the extent of their children's AGG, particularly during adolescence, given age-related increases in independence and ability to conceal problem behaviors (Burt, 2012). In addition, parent reports largely reflect behaviors occurring in the home and may not adequately capture AGG in other settings (e.g., school; Achenbach, McConaughy, & Howell, 1987). We thus computed mean scores based on data from all available informants (i.e., mother, teacher, and self-report) and examined them as the outcome in a series of sensitivity analyses. As shown in Tables S2.6 and S2.7 and Figure S2.2, results were fully consistent with those from the main analyses. In particular, the three-class LCGA solution yielded a "desisting" and a "persisting" trajectory group, with membership in the former predicted not only by low conflict and high income, but also by low ADI. In the main analyses, ADI was marginally significant as a predictor of desistance relative to persistence (p<.10). Additional studies are needed to elucidate the role of neighborhood advantage in predicting desistance from AGG. Regardless, our primary results do not appear to be specific to parent-reported AGG but instead persist to multi-informant composite reports as well.

Another methodological limitation concerns the nested structure of the data. As noted earlier, we could not include the CLUSTER command in the conditional LGM or in any of the person-centered models. We accounted for clustering in the conditional LGM by examining the data in “family-wide” format. As a verification, we also ran the conditional LGM with the data in “long” format, without controlling for nesting. Results were fully consistent with those reported in the main analyses (see Table S2.8). To verify results from the mixture models, in which we were unable to control for cluster effects, we ran the unconditional and conditional models with one twin per family randomly selected (see Table S2.9). Consistent with the results in the full sample, the three-class solution comprised a normative group following a consistently low trajectory and two other groups differentiated by their degrees of desistance. Results from the conditional model also replicated those in the full sample, with income and conflict distinguishing between desisting and persisting classes. We thus do not believe that our results were biased by the limitations of the software in accounting for clustering.

A limitation specific to the mixture models was that the average posterior probability was relatively low for the persisting trajectory. In mixture modeling, each participant is assigned to the group for which their posterior probability of membership is highest. Average posterior probabilities are thus a measure of classification accuracy, as higher values indicate greater certainty in class assignment. The average probability for the persisting group was .65, somewhat below the recommended threshold of .70 (Nagin, 2005). This is likely related to the study’s accelerated longitudinal design, with incomplete follow-up at the second and third waves of data collection. When the sample was restricted to those with complete data at all waves (N=370), average posterior probabilities for all three groups were > .85 (both the class solution and the percentage of the sample comprising each group were fully consistent with that observed in the full sample). In addition, average posterior probabilities for the desisting and low/stable trajectories were .76 and .92, respectively, in the full sample. The three-class solution thus assigned participants to trajectories with a reasonable degree of confidence, despite the planned missingness at later waves.

Lastly, all predictors of AGG trajectories were modeled as time-invariant based on scores at the middle childhood assessment, when sample size was largest. Although beyond the scope of the present study, changes in neighborhood conditions, neighborhood residence, household income and/or parenting behaviors may elicit changes in AGG. In the case of parenting, the association may be bidirectional (i.e., changes in AGG may evoke specific parenting behaviors, as suggested in prior work (Klahr, Klump, & Burt, 2014; Narusyte et al., 2011)). Future studies are needed to examine parallel trajectories of family characteristics and youth AGG, particularly to clarify how parenting and AGG develop in tandem.

Implications

Despite these limitations, the present study has several important implications. First, our findings underscore the importance of leveraging variable-centered *and* person-centered approaches when studying developmental trajectories, even when the trait in question appears to be dimensional. Variable-centered approaches are integral to understanding the overall, sample-level pattern as well as the predictors of the intercept variance. Given the heterogeneity in trajectories of psychopathology within typical samples, however, the magnitude and direction of change over time cannot be fully captured by a single slope estimate. Nor can we fully understand the predictors of the slope variance when collapsing across the entire sample. Thus, after determining via growth curve modeling that there is significant variance between persons in rate of change, one good solution would be to turn to person-centered techniques to understand the sources of that variance. By identifying discrete trajectories that differ not only by baseline level but also by direction/rate of change, person-centered techniques can both model the heterogeneity in participants' slopes and examine the predictors of this heterogeneity in a more nuanced way than can be done in variable-centered approaches.

In the present study, for example, the unconditional variable-centered analyses established that there was significant variance in participants' slopes, but the conditional analyses suggested that socioeconomic status did not significantly explain any of this variance, a finding that is highly inconsistent with prior research (Carroll, Mikhail, et al., 2023). Moreover, *low* parent-child conflict was implicated as a risk factor for continued AGG, despite predicting lower levels of AGG at baseline. It was only through the conditional mixture models that we were able to identify familial advantage and low parent-child conflict as specific predictors of desistance over time, and in a way that was consistent with theory and prior research. In short, the integration of the two approaches yielded a far more complete picture of the development of AGG, as well as a more logical conceptualization of the factors predicting interindividual variation, than either approach provided by itself.

Second, the present study advances our understanding of youth AGG as a developmental phenomenon unfolding within multiple contexts. Individual, familial, and neighborhood characteristics all emerged as important predictors of continuity and change in AGG. In particular, our findings implicate both the family and the neighborhood as critical contexts for the emergence of AGG, given that mother-child conflict, familial poverty, and neighborhood deprivation all contributed significantly to AGG trajectories. Moreover, among those youth with high levels of AGG at baseline, familial characteristics predicted which youth eventually desisted. The distinction between desistance and persistence among those with high levels of a given psychopathology is of great clinical importance, given that all who present for treatment presumably demonstrate elevated symptoms at baseline. Successful courses of treatment would thus be reflected in trajectories of desistance, meaning that factors promoting desistance are important to identify and leverage in targeted interventions.

The present findings specifically point to mother-child conflict as a potential treatment target for reducing youth AGG, consistent with the demonstrated efficacy and effectiveness of parent management training (PMT). PMT, a structured intervention in which parents are taught behavioral strategies to promote their children’s prosocial behavior and discourage conduct problems, has proven effective in reducing aggressive, disruptive, and noncompliant behaviors in children and adolescents (DeGarmo, Patterson, & Forgatch, 2004; DeGarmo & Forgatch, 2005; Hagen, Ogden, & Bjørnebekk, 2011; Kjøbli, Hukkelberg, & Ogden, 2013). Moreover, studies of treatment mechanisms have found increases in effective parenting (e.g., consistent discipline, positive involvement) to mediate the effects of PMT (DeGarmo et al., 2004; DeGarmo & Forgatch, 2005; Hagen et al., 2011). One goal for future work should be to determine to what extent parent-child conflict specifically serves as a mediator for the effects of parent training, as well as whether findings are consistent across informants (i.e., mother, father, and child reports of conflict).

Socioeconomic advantage within the home also appeared to interrupt trajectories of elevated AGG, consistent with prior work that has found exposure to deprivation to be a potent predictor of externalizing behaviors in particular (Carroll et al., 2023; Snyder, Young, & Hankin, 2019). These findings underscore the implications of socioeconomic disadvantage for youth mental health not only at a single timepoint but throughout the early developmental period. In more affluent households, for example, healthy development may be facilitated by greater access to resources, such as recreational facilities, structured activities, and, notably, mental health treatment. Indeed, the markers of desistance identified in the present study may partially reflect treatment effects, as youth from wealthier backgrounds who exhibit high levels of AGG may be better able to access treatment than their counterparts from disadvantaged homes. Future studies should examine access to mental health treatment as a potential mediator of the association between familial (dis)advantage and desistance from youth AGG. Studies are also needed to determine the extent to which changes in socioeconomic status over time predict trajectories of AGG.

Lastly, future research should employ variable and person-centered techniques to illuminate the effects of (dis)advantage on trajectories of other externalizing behaviors, such as non-aggressive rule-breaking (RB). In contrast to AGG, RB follows a developmental pattern characterized by increasing engagement during adolescence, as well as greater fluctuations and lower rank-order stability (Burt, 2012). One might thus expect to identify trajectories of RB that differ not by degree of desistance but rather by the extent of their *increase* across adolescence. Leveraging variable and person-centered methods would allow us to determine the risk factors that predict large increases in RB over time, which would represent promising targets for treatment efforts.

Tables

Table 2.1. *Descriptive statistics and correlations*

	1.	2.	3.	4.	5.	6.	7.
1. ADI	-						
2. Income	**-.40**	-					
3. Conflict[a]	.02	-.01	-				
4. Nurturance[a]	**-.15**	**.13**	**-.37**	-			
5. AGG T1[b]	**.15**	**-.16**	**.35**	**-.16**	-		
6. AGG T2[b]	**.08**	**-.11**	**.20**	**-.09**	**.39**	-	
7. AGG T3[b]	.04	-.06	**.17**	-.01	**.39**	**.62**	-
Mean (SD)	57.25 (22.68)	8.24 (2.68)	20.66 (4.63)	41.45 (3.53)	1.58 (1.87)	.69 (1.23)	.58 (1.13)
Range	2-99	1-10	12-39	22-48	0-8	0-8	0-6
N	2010	2010	2041	2041	2040	759	372

Note: Bold font indicates p<.05. [a]Mean scores were computed based on maternal and child reports. [b]Sum scores were computed based on maternal reports on eight CBCL items.

Table 2.2. *LGM fit statistics and parameter estimates*

Unconditional LGM model fit statistics					
Model	-2lnL	χ^2 (df)	AIC	BIC	SABIC
Linear growth	**5667.38**	**-**	**5679.38**	**5713.17**	**5694.10**
Means model	6119.52	452.14† (3)	6125.52	6142.41	6132.88

Conditional LGM parameter estimates			
Parameter	Estimate	S.E.	p-value
	Intercept		
Mean	**.692**	**.055**	**<.001**
Variance	**.311**	**.024**	**<.001**
ADI	**.055**	**.023**	**.017**
Income	**-.068**	**.025**	**.006**
Conflict	**.250**	**.019**	**<.001**
Nurturance	.017	.017	.325
Sex	**.210**	**.038**	**<.001**
Race	.016	.058	.784
	Slope		
Mean	**-.046**	**.008**	**<.001**
Variance	**.003**	**.001**	**<.001**
ADI	-.003	.004	.392
Income	.002	.004	.551
Conflict	**-.016**	**.003**	**<.001**
Nurturance	.001	.003	.849
Sex	**-.018**	**.006**	**.002**
Race	.005	.008	.548
	Covariances		
Intercept with slope	**-.022**	**.003**	**<.001**
Intercept 1 with Intercept 2			
MZ	**.274**	**.025**	**<.001**
DZ	**.163**	**.024**	**<.001**
Slope 1 with Slope 2			
MZ	**.003**	**.001**	**<.001**
DZ	**.001**	**.000**	**.002**
Intercept 1 with Slope 2			
MZ	**-.023**	**.003**	**<.001**
DZ	**-.013**	**.003**	**<.001**
Residual variance	**.135**	**.012**	**<.001**

Note: Bold font indicates $p<.05$. †Significant change in chi-square at $p<.05$.

Table 2.3. *LCGA fit statistics and predictors of group membership*

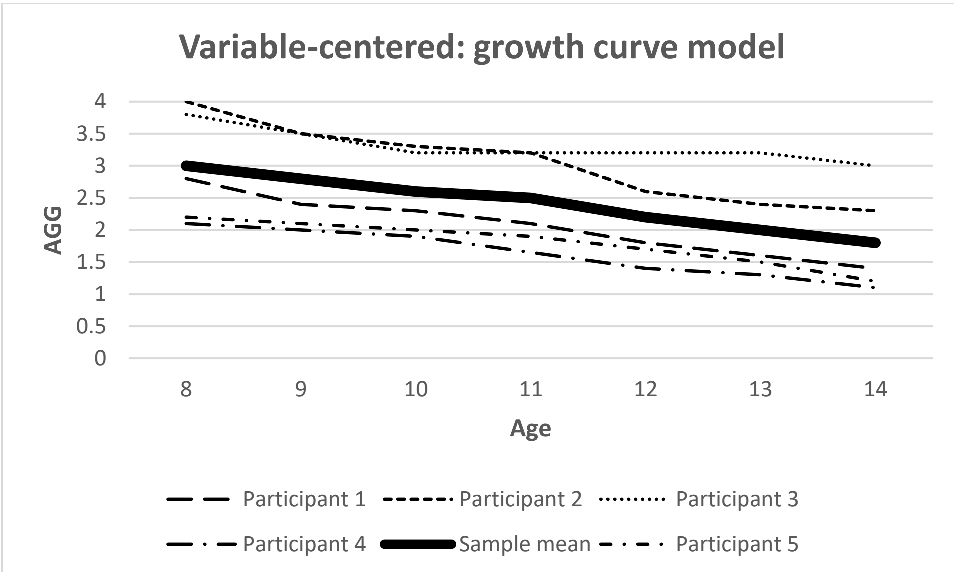
Unconditional LCGA model fit statistics			
	AIC	BIC	Smallest group (% of sample)
# of groups			
1	5983.49	6000.36	100.0
2	5467.89	5501.64	38.0
3	5197.92	5248.55	14.1
4	5077.70	5145.21	8.2
5	4987.81	5072.19	5.6
6	4931.54	5032.80	5.4
Conditional LCGA parameter estimates from 3-group model			
Persistent v. low/stable			
	Estimate	S.E.	*p*-value
ADI	**.310**	**.102**	**.002**
Income	**-.396**	**.086**	**<.001**
Conflict	**1.241**	**.112**	**<.001**
Nurturance	.147	.100	.144
Sex	**.717**	**.180**	**<.001**
Race	.250	.224	.265
Persistent v. desisting			
	Estimate	S.E.	*p*-value
ADI	.198	.119	.094
Income	**-.374**	**.095**	**<.001**
Conflict	**.369**	**.112**	**.001**
Nurturance	.128	.118	.277
Sex	.242	.218	.266
Race	.123	.259	.635
Desisting v. low/stable			
	Estimate	S.E.	*p*-value
ADI	.112	.078	.154
Income	-.022	.079	.780
Conflict	**.873**	**.084**	**<.001**
Nurturance	.019	.075	.805
Sex	**.475**	**.142**	**.001**
Race	.126	.185	.494

Note: Bold font indicates $p<.05$.

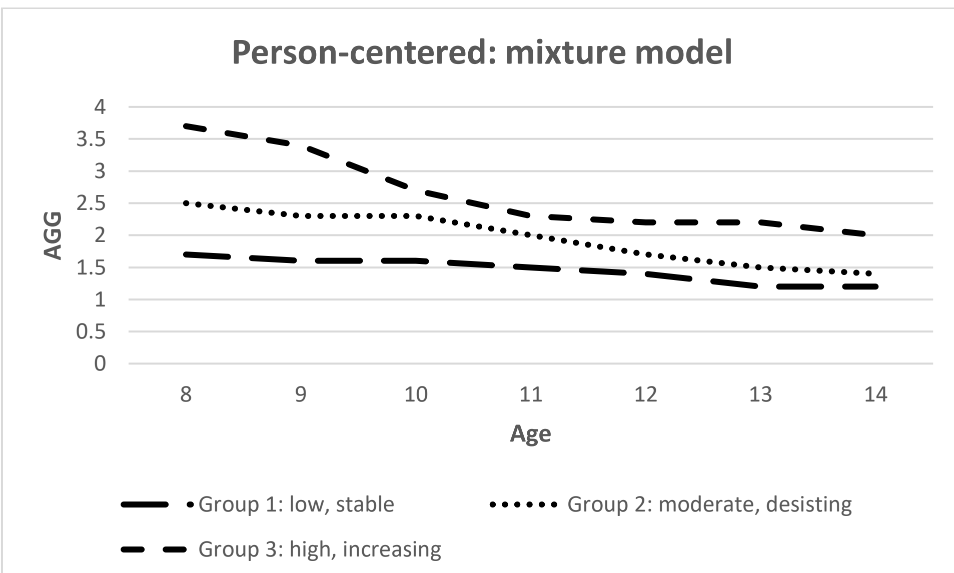
Table 2.4. *Characteristics of AGG trajectory groups*

	Class 1: low, stable	Class 2: desisting	Class 3: persisting
N	1079	681	290
% male	46.2	54.3	64.5
	Means (SD) [Range]		
ADI	55.62 (22.22) [2-99]	57.15 (23.28) [2-99]	63.62 (21.70) [5-99]
Income	8.42 (2.61) [1-10]	8.41 (2.52) [1-10]	7.22 (3.06) [1-10]
Conflict	19.23 (4.12) [12-34.5]	21.82 (4.40) [12.5-36.5]	23.34 (4.96) [12.5-39]
Nurturance	41.79 (3.49) [27-48]	41.23 (3.51) [22-48]	40.68 (3.61) [28-48]
AGG T1	.25 (.44) [0-2]	2.41 (1.24) [0-8]	4.56 (1.87) [0-8]
AGG T2	.24 (.58) [0-5]	.45 (.82) [0-6]	2.68 (1.54) [0-8]
AGG T3	.20 (.54) [0-3]	.34 (.71) [0-4]	2.48 (1.60) [0-6]
AGG T1 (log-transformed)	.17 (.30) [0-1.10]	1.17 (.34) [0-2.20]	1.65 (.40) [0-2.20]
AGG T2 (log-transformed)	.15 (.33) [0-1.79]	.26 (.42) [0-1.95]	1.22 (.40) [0-2.20]
AGG T3 (log-transformed)	.12 (.31) [0-1.39]	.21 (.38) [0-1.61]	1.13 (.51) [0-1.95]

Figures



a.



b.

Figure 2.1. *Variable-centered (a) and person-centered (b) approaches to modeling the development of AGG in a hypothetical childhood sample.* The variable-centered approach identified a mean-level trend across the entire sample, whereas the person-centered analysis extracted three distinct trajectory groups.

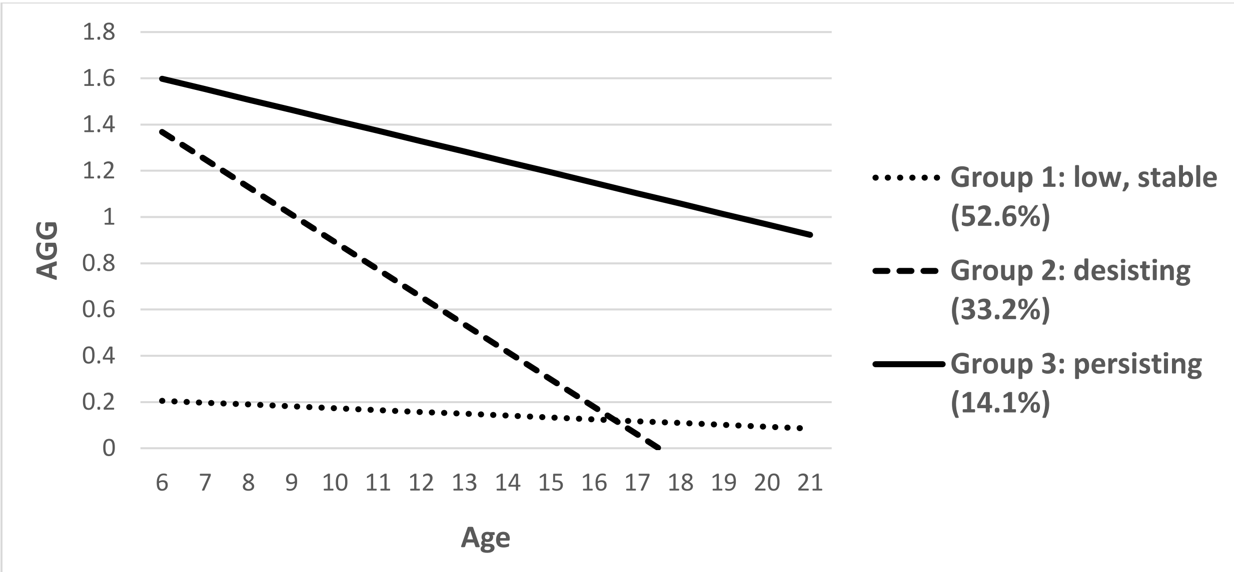


Figure 2.2. *Unconditional three-group model of AGG trajectories (person-centered).*

CHAPTER 3: UNDERSTANDING HETEROGENEITY IN TRAJECTORIES OF RULE-BREAKING VIA INTEGRATION OF VARIABLE-CENTERED AND PERSON-CENTERED METHODS

Abstract

Trajectories of antisocial behavior, including non-aggressive rule-breaking (RB), are characterized by considerable heterogeneity between persons. The best-known framework for conceptualizing this heterogeneity is Moffitt's person-centered taxonomy, which groups youth into qualitatively distinct trajectories based on age of onset. Although a large body of research supports Moffitt's distinction between life-course-persistent and adolescent-limited antisocial behavior, taxometric work has consistently indicated antisocial behavior to be a dimensional trait, without qualitative differences between persons. In other words, the theory implies that the field should use person-centered models, but the empirical data suggest antisocial behavior should be modeled using variable-centered statistical methods. Given this disconnect, it is surprising that nearly all longitudinal studies to date have employed only person-centered or variable-centered approaches, but not both. The present study sought to address this gap by examining the development of RB from middle childhood through to emerging adulthood via variable-centered and person-centered modeling, with socioeconomic disadvantage and peer delinquency considered as moderators. Both sets of analyses indicated that age-related increases in RB were normative, but variable-centered models were less useful for illuminating the association between socioeconomic disadvantage and the slope of RB. Across all analyses, peer delinquency emerged as a robust predictor of both persistence and escalation, consistent with Moffitt's hypotheses regarding the role of delinquent peer affiliation in trajectories of ASB. Overall, findings provided some additional support for Moffitt's taxonomy but also pointed to differences between trajectories that were a matter of degree, rather than kind, in line with the dimensional nature of youth behavioral problems.

Introduction

Antisocial behavior (ASB) comprises a broad range of actions that violate societal norms and/or the rights of other people. Behaviors range from relatively minor rule violations (e.g., lying, cheating) to more serious offenses that violate the law (e.g., robbery, assault). One defining characteristic of ASB is its continuity within-persons across early development. That is, the youth who exhibit the most behavioral problems in childhood often follow a pattern of persistent offending, culminating in a series of poor adult outcomes (Odgers et al., 2008). Despite this continuity, mean levels of ASB are known to change dramatically over time, with ASB observed to increase as much as tenfold in frequency during adolescence (Moffitt, 1993). The adolescent-increase in ASB is largely attributable to a spike in non-aggressive rule-breaking behaviors (RB), such as theft and vandalism, which are rare during childhood but increase in both prevalence and incidence during the teen years. These developmental patterns are consistent with the traditional conceptualization of the age-crime curve, one of the most robust findings in the field of criminology, with more youth observed to engage in delinquent activities during adolescence than at any other life stage (Hirschi & Gottfredson, 1983).

In what is now a classic taxonomy, Moffitt (1993) proposed the peak of the age-crime curve to represent the activities of two qualitatively distinct groups of youth: life-course-persistent offenders and adolescent-limited offenders. Life-course-persistent offenders were hypothesized to exhibit a broad spectrum of problem behaviors, often beginning with physical aggression (AGG) during the preschool years and escalating to include more severe aggressive and non-aggressive offenses throughout childhood and adolescence. Adolescent-limited offenders, by contrast, represented typically developing youth whose engagement in (primarily non-aggressive) ASB was transient and largely attributable to social factors, such as affiliation with delinquent peers. The remaining youth were thought to abstain from ASB throughout childhood and adolescence.

One essential feature of Moffitt's framework is its qualitative approach, with youth classified into subgroups based on shared characteristics (e.g., age of onset). While youth *within* a particular subgroup are hypothesized to generally follow the same trajectory of ASB, trajectories are thought to differ *between* subgroups not only by degree but also by kind (e.g., escalating versus desisting). Each subgroup is also thought to be characterized by distinct correlates, with life-course-persistent ASB linked to more childhood risk factors and poorer life outcomes than all other trajectories. In Moffitt's original framework, however, there was not a one-to-one correlation between the groups' degree and kind of engagement in ASB and their psychological functioning. Specifically, Moffitt posited that youth who demonstrated adolescent-limited ASB were in fact better adjusted than those who abstained completely, as the latter were believed to exhibit personality traits (e.g., behavioral inhibition, poor social skills) that left them excluded from adolescent peer networks (Moffitt, 1993).

Since the original publication, Moffitt's taxonomy has been updated in two key ways (Moffitt, 2003). First, the taxonomy expanded to include a childhood-limited group of youth, who largely desisted from ASB by adolescence. Second, the framework shifted to a more continuous conceptualization of the groups' relative psychological functioning. That is, the abstainers were viewed as the healthiest group, followed by the childhood-limited, adolescent-limited, and life-course-persistent groups, respectively. Follow-up data in adulthood from the Dunedin Longitudinal Study indicated that those youth who abstained from ASB had the best outcomes across domains, whereas adolescent-onset youth were second only to those with life-course-persistent trajectories in psychological, financial, and legal problems (Moffitt, Caspi, Harrington, & Milne, 2002). In other words, adolescent-onset ASB was both relatively normative and linked to a series of poor outcomes in early adulthood.

Consistent with the relatively normative nature of adolescent engagement in ASB, person-centered studies spanning adolescence have often reported a modal trajectory of increasing RB (e.g., Diamantopoulou, Verhulst, & van der Ende, 2011; Mata, 2013). Person-centered studies have largely used a method developed by Nagin, termed semi-parametric mixture modeling (Nagin, 1999). In this approach, participants are assigned to latent, or unobserved, trajectory groups based on their posterior probability of group membership (i.e., their likelihood of belonging to a particular group given their scores on the outcome measure at each timepoint). Although groups identified via semi-parametric mixture modeling are free to differ from one another by intercept (i.e., level of the outcome measure, often at baseline) and slope (i.e., rate of change), participants within the same group are assumed to share the same trajectory (see Figure 3.1a).

Using these methods, Diamantopoulou and colleagues found that even “low” trajectory groups, which comprised the majority of participants, exhibited increases in RB between ages 11 and 18. Likewise, nearly half of participants in the Montreal Longitudinal-Experimental Study followed a trajectory of moderate/increasing non-violent offending during adolescence, while a much smaller percentage of the sample (<10%) followed a chronic trajectory (Fontaine, Lacourse, Vitaro, & Tremblay, 2014). Consistent with Moffitt’s updated theory, both the “moderate” and “chronic” trajectory groups demonstrated poorer adult outcomes than did the “low” group. Other person-centered studies have similarly found many, if not most, youth to exhibit time-limited increases in RB but relatively few to demonstrate chronic engagement (Barker et al., 2007; van Lier et al., 2009), as Moffitt hypothesized.

Despite the support of Moffitt’s theories in person-centered analyses, it is worth noting that some studies of ASB report variability *within* groups in baseline level and rate of change (e.g., Gardner, 2006; Schaeffer et al., 2003, 2006). These studies make use of growth mixture modeling (GMM), a more flexible alternative to Nagin’s approach that allows for both between and within-class variation (Muthén & Muthén, 2000). Although relatively few studies have leveraged GMM, these findings nevertheless reflect a deviation from Moffitt’s hypotheses. Consistent with this, other work has strongly indicated that person-centered methods do not accurately represent the underlying structure of ASB. Specifically, taxometric analyses have found data on youth ASB to conform to simulated dimensional data but not to simulated categorical data, and these findings persisted to both cross-sectional and longitudinal analyses (Walters, 2011; Walters & Ruscio, 2013). Indeed, Burt (2012) argued that Moffitt's taxonomy could be re-conceptualized in dimensional terms, with childhood-onset and adolescent-limited ASB linked to AGG and RB, respectively. Because engagement in ASB thus varies continuously across the population, some researchers argue that it has no qualitatively distinct underlying classes and is best-represented using variable-centered models (e.g., Walters & Ruscio, 2013).

In longitudinal variable-centered approaches, growth curve models are used to quantify the mean change over time in the trait of interest across the entire sample (see Figure 3.1b; McArdle & Epstein, 1987). As in person-centered approaches, growth curve models capture individual patterns of continuity and change by the intercept and slope growth factors. Rather than considering the sample to comprise distinct subgroups, however, growth curve models treat all participants as belonging to the same underlying population, with patterns of growth that vary continuously from one another. The means of the intercept and slope represent the typical pattern of growth across the entire sample, whereas the variances of the intercept and slope represent individual differences in the baseline level of the outcome and the rate of change, respectively (Preacher, Wichman, MacCallum, & Briggs, 2008). As an example, a variable-centered analysis of RB might find the mean linear slope to be positive, and the intercept and slope variances to each be significant. These findings would suggest that the average participant exhibited an increase in RB over time, but that there were significant between-person differences in both initial level and rate of change.

In some ways, variable-centered studies of RB have yielded results consistent with those from the person-centered literature. For example, the sample-wide trend is typically one of increasing RB during adolescence (e.g., Doornwaard, Branje, Meeus, & ter Bogt, 2012; Padilla-Walker, Memmott-Elison, & Coyne, 2018), followed by declining or plateauing engagement by early adulthood (e.g., Doornwaard et al., 2012). Moreover, most studies have reported significant variability between persons in both the intercept and slope of RB (Carroll, Mikhail, &

Burt, 2023), as well as significant differences in baseline engagement as a function of biological sex (e.g., Yan, Schoppe-Sullivan, & Beauchaine, 2021), neighborhood disadvantage (Ingoldsby, 2003; Thaweekoon, 2006), and exposure to delinquent peers (Ingoldsby, 2003).

Nevertheless, and despite their ability to correctly model the actual development of RB as a dimensional trait, variable-centered methods have limited potential to capture the heterogeneity observed in most samples. Even if the mean slope indicates a small age-related increase in RB, for example, a small percentage of participants may decline in their engagement over time, and others may exhibit a large increase. The primary strength of person-centered methods lies in their ability to identify these subgroups of youth, whose differing patterns of engagement cannot be fully captured in a single slope estimate. Indeed, proponents of person-centered approaches contend that the groups serve as an approximation of participants' trajectories and need not exist in reality to be theoretically and empirically useful (Nagin & Tremblay, 2005). For instance, person-centered analyses could facilitate the identification of risk factors that distinguish between youth who exhibit large adolescent-increases in RB and their counterparts whose engagement falls within normal limits, a finding with clear implications for intervention work.

The considerations discussed thus far raise a key question for developmental researchers, namely, how can trajectories of ASB be modeled in a way that takes into account both developmental theory and the underlying structure of ASB data? That is, the most widely accepted developmental theory is Moffitt's person-centered taxonomy, which provides a useful framework for understanding the heterogeneity within ASB but does not accurately represent its dimensional structure. One way to resolve this apparent disconnect is to leverage both person-centered and variable-centered statistical approaches in the same dataset, which allows for a far more complete and accurate understanding of development than either can provide by itself.

A recent review (Carroll, Mikhail, et al., 2023) began the process of integrating these two approaches through a conceptual review of the developmental literatures on antisocial behavior.

Of the 48 studies of RB included in the review, only one (Ingoldsby, 2003) leveraged both statistical techniques, with largely consistent results across methods. For instance, the modal trajectory was one of moderate RB that declined over time, and the growth curve model also indicated a mean-level decline. Both sets of analyses also identified exposure to neighborhood disadvantage and delinquent peers as predictors of RB trajectories. However, participants were only followed until age 10, meaning that neighborhood and peer influences were not examined as predictors of RB trajectories during adolescence. This represents a potentially important gap in the literature, in light of studies indicating that the extent of delinquent peer affiliation increases for most youth from childhood to adolescence (Wang & Dishion, 2012; Yoon, Yoon, Yoon, & Snyder, 2019). Neighborhood and peer influences may also be more salient to behavioral outcomes in adolescence relative to childhood, given age-related increases in mobility and independence (Leventhal & Brooks-Gunn, 2000). In short, because no study of RB to date has leveraged both person-centered and variable-centered approaches together in a sample spanning adolescence, the factors giving rise to the varied patterns of RB (i.e., desistance, persistence, and escalation) typically observed throughout early development are not well understood.

Present Study

The aim of the present study was to illuminate the development of RB from middle childhood into emerging adulthood in a large, population-based sample enriched for neighborhood disadvantage. Trajectories of RB were modeled via a series of variable-centered and person-centered analyses, with biological sex, race/ethnicity, socioeconomic disadvantage, and peer delinquency examined as potential moderators. In light of theoretical work that views youth development as embedded within multiple contexts (Bronfenbrenner, 1988) and empirical studies indicating that familial and neighborhood disadvantage each contribute uniquely to youth outcomes (Carroll, et al., 2023; Kupersmidt et al., 1995), we examined both household income and a composite measure of neighborhood deprivation as moderators. Consistent with prior work (e.g., Doornwaard et al., 2012; Isen et al., 2022), we hypothesized that 1) the mean level of RB would increase across the study period, 2) there would be either three or four RB trajectory groups, and 3) male sex, lower household income, greater neighborhood disadvantage, and greater peer delinquency would predict higher levels of RB at baseline as well as persistent and escalating engagement across adolescence.

Methods

Transparency and Openness

The present study was not preregistered. Because of the language in the informed consent document at intake, we cannot post the data publicly, but they can be obtained from the primary author upon reasonable request. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All procedures were approved by the primary author's institutional review board and are in compliance with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration. Child participants provided informed assent, and parents provided informed consent for themselves and their children.

Participants

The population-based Michigan State University Twin Registry (MSUTR) includes several independent twin projects. Participants in the current study were drawn from the Twin Study of Behavioral and Emotional Development in Children (TBED-C), a study within the MSUTR. The TBED-C includes a population-based arm (*N*=528 families), and an 'under-resourced' arm for which inclusion criteria also specified that participating families lived in neighborhoods with neighborhood poverty levels at or above the Census mean at study onset (10.5%) (*N*=502 families). To recruit families at intake, the Department of Vital Records in the Michigan Department of Health and Human Services (MDHHS) identified twins in our age-range via the Michigan Twins Project (MTP), a population-based registry of more than 32,000 twins in Lower Michigan recruited via birth records. The Michigan Bureau of Integration, Information, and

Planning Services database was used to locate family addresses no more than 120 miles from East Lansing, Michigan through parent drivers' license information. Pre-made recruitment packets were then mailed to parents by the MDHHS. Parents who did not respond to the first mailing were sent additional mailings roughly one month apart until either a reply was received or up to four letters had been mailed.

This recruitment strategy for the TBED-C yielded an overall response rate of 57% for the under-resourced arm and 63% for the population-based arm. Other recruitment and sampling details can be found in prior publications (i.e., Burt & Klump, 2019). The two arms of the study were analyzed jointly for the current analyses and were 48.7% female and 51.3% male. Across the full sample, participants endorsed racial/ethnic identities in the following proportions: white (non-Latinx), 82%; Black, 10%; Latinx, 1%; Asian, 1%; Indigenous, 1%; multiracial, 6%. However, families in the under-resourced arm, but not the population-based arm, were more racially diverse than the local population (e.g., 14% Black and 77% white in the under-resourced arm versus 5% Black and 87% white in the population-based arm; 5% Black and 85% white in the local area Census).

Because 90+% of TBED-C families were recruited out of the MTP, we were able to use the MTP data to compare families who chose to participate in TBED-C with those who were recruited but did not participate. TBED-C families were generally representative of recruited but non-participating families. As compared to non-participating twins, participating twins reported similar levels of conduct problems, emotional symptoms, and hyperactivity (*d* ranged from -.08 to .01 in the population-based arm and .01 to .09 in the under-resourced arm; all ns). Participating families also did not differ from non-participating families in paternal felony convictions (*d* = -.01 and .13 for the population-based and under-resourced arms, respectively), rate of single-parent homes (*d* = .10 and -.01 for the population-based and under-resourced arms, respectively), paternal years of education (both $d \leq .12$), or maternal and paternal alcohol problems (*d* ranged from .03 to .05 across the two arms). However, participating mothers in both samples reported slightly more years of education (*d* = .17 and .26, both *p* < .05) than non-participating mothers. Maternal felony convictions differed across participating and non-participating families in the population-based arm (*d* = -.20; *p* < .05) but not in the under-resourced arm (*d* = .02). In short, our recruitment procedures appear to have yielded a sample that was representative of both recruited families and the general population of the State of Michigan.

Behavioral data regarding the children's RB were collected at up to three time points. All 2,060 twins in 1,030 families were assessed once in middle childhood (ages 6-11; M_{age} (SD)=8.02 (1.49)) as part of the TBED-C (Wave 1). TBED-C participants residing in modestly-to-severely disadvantaged neighborhoods (*N* = 768 families) are currently being recruited for reassessment as adolescents up to two times, 18-months apart, through the Michigan Twin Neurogenetics Study (MTwiNS). For the present study, the sample sizes at Waves 2 and 3 were 760 and 414, respectively.

Procedure

TBED-C families completed their intake assessments between 2008 and 2015. Assessment teams consisted of two research assistants and at least one paid staff member and took 4-5 hours to complete (lunch was provided). Families completed questionnaires, interviews, and videotaped interactions during the in-person assessment. The primary caregiver (nearly always the mother) participated alongside the twins. Assessments typically took place in our university laboratory (903 families). If families were unable or unwilling to travel, however, assessments took place in participants' homes (13%). Families with older twins were less likely to complete home visits than on-site assessment visits (Cohen's *d*=-.24, *p*=.01), as were families that identified as White (*d*=-.21; *p*<.05). Similarly, families with fewer financial resources were more likely to participate in home visits versus on-site assessments (*d*=-.32; *p*=.004). However, families completing on-site assessments versus home visits did not differ in maternal education, zygosity of the twins, or sex of the twins (*d*s ranged from -.02 to .16, all ns). The

MTwiNS follow-up assessments were also completed in-person. The twins and their primary caregiver completed an in-person assessment lasting 4-8 hours at either the East Lansing or Ann Arbor-based laboratories.

Measures

RB

Youth RB was defined via a combination of caregiver, teacher, and youth reports. The primary caregiver (nearly always the twins' mother) completed the 17-item RB scale from the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001) at all three waves, rating the extent to which a series of statements described the child's behavior over the past six months using a three-point scale (0=never to 2=often/mostly true). Caregiver-informant reports were available for 99% of participating twins at Wave 1, 99% at Wave 2, and 90% at Wave 3. Teacher reports of RB (12 items) were obtained at the first two assessment waves via the Achenbach Teacher Report Form (TRF; Achenbach & Rescorla, 2001). Teachers rated the twins' behaviors during the preceding six months using the three-point scale described above. The teachers of 115 participants were not available for assessment (because the twins were home-schooled or because parental consents to contact the teachers were completed incorrectly, etc.). Our teacher participation rate across both subsamples was 86% at Wave 1 and 60% at Wave 2, with teacher reports available for 1,551 and 453 participants, respectively. In addition, the twins completed the Youth Self-Report (YSR) at Waves 2 and 3, reporting on their own RB (15 items) during the preceding six months using the three-point scale described above (Achenbach & Rescorla, 2001). Reports were available from 99% and 83% of the twins at Waves 2 and 3, respectively.

The RB scale from the Achenbach family of instruments assesses a wide range of behaviors, including minor rule violations (e.g., lying, cheating) as well as illegal behaviors (e.g., vandalism, setting fires). The latter had extremely low base rates in our sample, particularly at the childhood assessment. We thus constructed a more streamlined scale of RB items, reflecting behaviors that were present to some degree at all assessment waves, via a series of item response theory (IRT) analyses. We also conducted a series of measurement invariance analyses to confirm that the items retained in the IRT models assessed RB consistently across measurement waves (see Supplementary Methods in Appendix C).

Familial Socioeconomic Disadvantage

Familial socioeconomic disadvantage was assessed via maternal reports of annual household income at all assessment waves. Household income was measured on a 10-point Likert scale (1 = <$10,000 to 10 = >$50,000) at Wave 1, and on a 13-point Likert scale (0 = <$5,000 to 12 = >$90,000) at Waves 2 and 3. To account for the different scales, we constructed a new eight-point scale (1 = <$10,000 to 8 = >$50,000) (e.g., families reporting an income <$5,000 at Wave 2 were coded as belonging to the <$10,000 bracket). Mean composite scores were computed to represent average income across all three waves.

Neighborhood Socioeconomic Disadvantage

Neighborhood disadvantage was defined via the Area Deprivation Index (ADI; Kind & Buckingham, 2018), which comprises 17 indices of Census block group disadvantage (e.g., poverty rate, income disparity). Data were weighted according to the factor loadings identified by Kind & Buckingham (2018), and weighted variables were summed to create a deprivation index score for each Census block group. Families were assigned a percentile indicating the level of deprivation in their block group relative to that of all U.S. block groups, with higher scores indicating greater deprivation. Mean ADI scores were 57.24 (SD=22.67; range: 2-99), 61.57 (SD=20.94; range: 6-99), and 60.71 (SD=21.62; range: 12-99) at the TBED-C and MTwiNS assessments, respectively. These scores correspond to a moderate level of neighborhood disadvantage, albeit with considerable variation across participating families. At the TBED-C assessment, for instance, Census tract poverty rates ranged from 0 to 93%, with a mean of 19%. Mean ADI scores were computed to represent the average level of neighborhood deprivation over time.

Peer Delinquency

Primary caregivers reported on their children's peer group affiliations at Waves 1 and 2 using the five-item Peer Delinquency scale from the Friends Inventory (Walden, McGue, Burt, & Elkins, 2004). Caregivers were instructed to provide ratings for each child's entire peer group, with items scored on a four-point scale (1 = "none of my child's friends are like that" to 4 = "all of my child's friends are like that"; e.g., "my child's friends break the rules"). Caregiver reports were available for 96.4% and >99% of participants at Waves 1 and 2, respectively. The twins also reported on their own peer groups' delinquency at Waves 2 and 3 using the Friends Inventory. Items are nearly identical across parent and child versions of the Friends Inventory, with minor alterations in wording. Self-reports were available for >99% of participants at both waves. Parent and child reports were averaged across waves to represent mean peer delinquency over time.

Race/Ethnicity

Given inequalities in exposure to socioeconomic disadvantage across racial/ethnic identities (Peterson & Krivo, 2009), race was included as a covariate in all longitudinal analyses, coded dichotomously as white/racially marginalized due to the composition of our sample.

Data Analytic Strategy

Multi-informant mean RB scores were computed for each assessment wave based on the items retained in the final IRT model (see Supplementary Methods and Results in Appendix C for additional details, including confirmation of measurement invariance). Mean RB scores were log-transformed to adjust for positive skew and were subsequently examined as the outcome measures in the variable-centered and person-centered analyses. Continuous covariates (i.e., income, neighborhood disadvantage, and peer delinquency) were standardized for ease of interpretation.

Variable-Centered Analyses

Analyses were conducted in M*plus* 8.4 (Muthén & Muthén, 1998-2019) using a robust maximum likelihood estimator and full information maximum likelihood estimation (FIML) to account for missing data on the outcome. We examined phenotypic changes in the frequency of RB over time across the full sample via latent growth curve modeling, which quantifies within-person changes in the outcome of interest as well as between-person variation in these changes (McArdle & Epstein, 1987). First, unconditional models were run to determine whether the frequency of RB increased, decreased, or remained constant with age. Age was examined in lieu of assessment wave as the metric of time given the wide range of ages included at each assessment (see Table S3.1). The definition variable approach was used to scale time in years, with the factor loadings constrained to vary by age (Sterba, 2014). Because the definition variables must be complete for the model to be estimated correctly, placeholder values were imputed to represent age for participants with missing data (Grimm, Ram, & Estabrook, 2017). Specifically, the mean number of years that elapsed between waves was added to participants' ages at Wave 1 if data were missing at Waves 2 and/or 3. Age was available for all participants at Wave 1. Imputed values had no impact on estimation, as missing data on the definition variables was connected with missing data on the outcomes.

Intercepts were centered to represent RB at age six, the youngest age at the first TBED-C assessment wave. The slope factor represented annual change. Because participants were assessed on only three occasions, we did not examine non-linear growth. Initial analyses compared the respective fits of an intercept-only ("no-growth") model and a model allowing for linear growth. The former had three parameters (intercept mean, intercept variance, and residual variance), and the latter had six (intercept and slope means, intercept and slope variances and their covariance, and residual variance). Model fit was evaluated with four likelihood-based indices: the chi-square difference test (Bollen, 1989), Akaike information criterion (AIC; Akaike, 1987), Bayesian information criterion (BIC; Raftery, 1995), and sample-size adjusted Bayesian information criterion (SABIC; Sclove, 1987). For all indices, lower values indicated better model fit. In the unconditional models, the CLUSTER command was used to account for the nesting of twins within families.

The best-fitting model was selected for the conditional growth curve analyses. RB was fit conditional upon sex, race, income, ADI, and peer delinquency scores, all of which were averaged across waves and subsequently examined as time-invariant predictors of the intercept and slope. Because M*plus* employs listwise deletion for missing covariates, all available data were analyzed in the conditional growth curve model by using multiple imputation (Little & Rubin, 2019), with five imputed datasets. As a result, estimates of the parameters and standard errors reflect the uncertainty that is due to missing data. In M*plus*, it is not possible to include the CLUSTER command while using multiple imputation. Thus, in the conditional LGM, the data were structured to be in "family-wide" format, with parameters constrained to equality across both twins in a pair. The conditional LGM also accounted for the difference in within-pair resemblance for monozygotic (identical) and dizygotic (fraternal) twins, as identical twins are expected to resemble one another more so than fraternal twins on virtually every trait (Turkheimer, 2000).

Person-Centered Analyses

Latent class growth analyses were conducted in M*plus* 8.4 using a robust maximum likelihood estimator and FIML to account for missing outcome data. We examined participants' growth in RB as a function of their age at each assessment wave. We began with a one-class model and progressed through to a six-class model, consistent with the steps outlined by Muthén (2004) and with the approach employed in prior studies examining trajectories of ASB (e.g., Odgers et al., 2008). All models allowed the groups to differ from one another by intercept and slope but constrained within-group intercept and slope variances to zero. The optimal model was determined via consideration of two fit indices, the AIC and BIC, as well as the interpretability of the class solution (i.e., the percentage of the sample assigned to each class). Biological sex, race, income, ADI, and peer delinquency were subsequently examined as time-invariant predictors of class membership, using five imputed datasets to account for missing values. Notably, it is not possible to account for clustering when examining age as the metric of time in M*plus* mixture models, nor was it feasible to examine the data in “family-wide” format because both twins in a pair would be assigned to the same trajectory by default. Thus, we did not account for clustering in our main analyses but ran supplemental analyses with one twin randomly selected per family, discussed in greater detail later in the manuscript.

Results

Descriptives

Descriptive statistics and correlations across assessment waves are shown in Table 3.1, and descriptives by age are shown in Table S3.1. Participants evidenced moderate rank-order stability in RB over time. Paired-sample *t* tests indicated that mean levels of RB did not change significantly across assessment waves, although the increase from Wave 2 to Wave 3 was marginally significant ($t(373) = 1.894$, $p = .059$). At Waves 1 and 2, males had higher levels of RB than females (Cohen’s *d*s were .36 and .28, respectively, both *p*s <.001); sex differences were not significant at Wave 3. In addition, ADI and income were significantly correlated with RB at the first two time points, but not at the third. Peer delinquency was significantly correlated with RB at all waves. All correlations were in the expected direction.

Variable-Centered

Unconditional models

Initial analyses compared the respective fits of an intercept-only model and a linear growth model to test the hypothesis that engagement in RB increased with age. The linear growth model fit better than the intercept-only model according to all fit indices (see Table 3.2) and was thus selected as the final unconditional model. At age six, the average RB score was .832 (SE = .029, $p< .001$), and scores increased by .018 points per year on average (SE = .005, $p< .001$), consistent with hypotheses. There was significant variation between-persons in both baseline level ($b = .538$, SE = .051, $p< .001$) and rate of change ($b = .005$, SE = .001, $p< .001$). The estimated covariance between the intercept and the slope was -.038 (SE = .006, $p< .001$), indicating that participants with higher levels of RB at baseline exhibited somewhat less of an increase with age.

Conditional model

We subsequently examined ADI, income, peer delinquency, sex, and race as predictors of continuity and change in RB. As hypothesized, lower income, greater peer delinquency, and male sex predicted higher levels of RB at baseline. Income, peer delinquency, and sex also predicted rate of change (see Table 3.2). Specifically, males increased somewhat less in their RB over time relative to females, whereas participants with higher levels of peer delinquency exhibited a greater increase. Unexpectedly, higher income also predicted a greater increase in RB over time. Neither ADI nor race predicted the intercept or the slope.

Person-Centered

Unconditional models

Initial analyses tested the hypothesis that there would be either three or four distinct RB trajectories. Model fit improved as the number of classes increased (see Table 3.3); however, the five and six-class models each contained groups that were not meaningfully distinct from one another. For example, the five-class model contained two groups with high, persistent RB. Based on both fit and theoretical considerations, the four-class model was selected as the optimal solution. As shown in Figure 3.2, the four-class model comprised two trajectories with low RB at baseline and two with high RB at baseline. The former were distinguished by the extent of their increase over time; one group, which comprised nearly 40% of the sample, exhibited a small increase with age, while the other, smaller group escalated substantially. Of the groups with high RB at baseline, one fully desisted by the end of adolescence, whereas the other followed a chronic trajectory.

Conditional model

We first examined the predictors of membership in the chronic trajectory relative to the normative group. These groups differed as a function of ADI, income, peer delinquency, and sex. Specifically, greater neighborhood deprivation, lower income, greater affiliation with delinquent peers, and male sex predicted chronic RB trajectories. To elucidate the factors predicting the slope, we subsequently examined each pair of groups with similar RB at baseline. ADI and peer delinquency distinguished between the escalating and normative groups, such that youth who resided in disadvantaged neighborhoods and associated with more delinquent peers were more likely to increase substantially in their RB. High levels of peer delinquency also predicted membership in the chronic trajectory relative to the desisting trajectory, as did low income. Neither sex nor race distinguished between either pair of trajectory groups (see Tables 3.3 and S3.6 for conditional model results and Table 3.4 for trajectory group characteristics).

Interpretive Integration

Variable-centered and person-centered analyses both indicated that an increase in RB over time was normative. Moreover, lower household income, greater peer delinquency, and male sex all predicted elevated RB at baseline, as they predicted the intercept of the growth curve and distinguished between the chronic and normative trajectory groups. Regarding predictors of the slope, however, results were consistent across modeling approaches for some but not others. Both sets of analyses identified peer delinquency as a critical predictor of escalating engagement. That is, greater peer delinquency predicted a greater age-related increase in RB in the growth curve model, as well as membership in the escalating and chronic groups relative to the normative and desisting groups, respectively, in the mixture models. However, familial affluence predicted a *greater* increase in RB in the growth curve, a finding that may reflect a statistical artifact, as those from wealthier families had lower levels of RB at baseline and thus had more room to increase over time. Person-centered results, by contrast, indicated that income distinguished between the chronic and desisting groups, and in the expected direction. Neighborhood deprivation also distinguished between the escalating and normative trajectories but did not predict the slope of the growth curve, possibly because the escalating trajectory comprised a small portion of the sample. Such findings suggest that variable-centered approaches are limited in their ability to elucidate the factors predicting membership in small, yet clinically meaningful, trajectory groups.

Discussion

The present study aimed to elucidate trajectories of RB across early development by integrating results from person-centered and variable-centered modeling approaches. Consistent with our hypotheses, age-related increases in RB were normative. That is, the sample-wide pattern identified in the variable-centered analyses was one of increasing RB across the study period, and person-centered analyses yielded a modal trajectory characterized by a small increase. Taken together, these findings are consistent with the broader empirical literature, which has reported that most (but not all) youth follow trajectories of increasing RB during adolescence (e.g., de Haan, Prinzie, & Dekovic, 2012; Diamantopoulou et al., 2011). The person-centered results are also largely consistent with Moffitt’s taxonomy (Moffitt, 1993, 2003), with a subgroup of youth exhibiting a life-course-persistent trajectory and a larger portion of the sample demonstrating onset in adolescence.

Nevertheless, the development of RB was far from uniform across individuals and trajectory groups. For instance, nearly half (46.6%) of the sample followed a trajectory consistent with Moffitt’s adolescent-limited class, with little-to-no engagement during childhood, but the two adolescent-limited groups identified here differed sharply in the magnitude of their age-related increase. The larger of the two groups exhibited a modest increase, consistent with the mean pattern identified in the growth curve, whereas the other group, which comprised 15.5% of adolescent-increasers, eventually surpassed the chronic trajectory in level of engagement. The heterogeneous patterns of change identified here raise a question that person-centered approaches are uniquely positioned to answer, namely, *what factors predict escalating RB among youth with few behavioral problems in childhood?*

The mixture models found ADI and peer delinquency to predict membership in the escalating trajectory relative to the normative trajectory. Specifically, those who escalated were more likely to reside in disadvantaged neighborhoods and affiliate with delinquent peers, consistent with theory (Bronfenbrenner, 1988; Moffitt, 1993, 2003) and prior research (Carroll et al., 2023; Ettekal & Ladd, 2015). Higher levels of peer delinquency also predicted a greater age-related increase in RB in the growth curve model. Results were not consistent across modeling approaches for ADI, however, as ADI did not predict the slope of the growth curve. These discrepant findings may be due to the size of the escalating trajectory, which comprised only ~7% of participants and was likely not well-characterized via analyses that collapsed across the entire sample. Person-centered models were thus better positioned to illuminate the role of neighborhood disadvantage in predicting escalating RB, although both sets of analyses implicated peer delinquency as an important predictor of the slope.

Integration of the person-centered and variable-centered results also yields a fuller understanding of the factors predicting desistance from RB. Lower levels of peer delinquency predicted membership in the desisting trajectory relative to the chronic one, consistent with the results for peer delinquency in the growth curve model. For income, however, the growth curve indicated that affluence was associated with a *greater* increase in RB with age, a finding that may reflect regression to the mean, as suggested by prior variable-centered simulations examining predictors of change (Marsh & Hau, 2002). In the person-centered analyses, by contrast, higher income predicted desistance relative to persistence. Thus, while the growth curve model accurately identified the association between income and RB at baseline, consideration of results from the mixture model was necessary to understand the association between income and rate of change.

The present study also found that ADI and income incremented one another as predictors of membership in the chronic trajectory group relative to the normative one, consistent with prior work indicating that exposure to disadvantage in multiple contexts is particularly salient to the development of externalizing (Carroll et al., 2023; Kupersmidt et al., 1995). Not only did disadvantage predict elevated RB at baseline in our sample in both sets of analyses, but it also predicted escalation relative to normative engagement and persistence relative to desistance in the mixture models. These findings underscore the long-lasting implications of socioeconomic disadvantage for youth mental health. Youth who reside in disadvantaged neighborhoods may face economic insecurity, household overcrowding, and/or community violence, among other risk factors (Singh, 2003). Additional research is needed to determine which specific aspects of neighborhood disadvantage serve as 'active ingredients' for trajectories of persistent or escalating RB. In turn, our finding that familial affluence predicted desistance may partially reflect treatment effects, as youth from wealthier backgrounds who demonstrate childhood-onset RB may be more likely to access treatment than their less affluent peers. Future studies should examine whether access to psychological treatment explains the association between household income and recovery from RB.

The variable-centered and person-centered results together also shed light on sex differences in RB. The former indicated that males exhibited greater engagement at baseline, consistent with the broader variable-centered literature on RB, which has typically found sex to predict the intercept (Carroll, Mikhail, et al., 2023). Male sex also predicted a somewhat smaller increase with age in the growth curve, a finding that is less consistent with prior research. Results from the mixture models help to clarify the association between sex and rate of change. Male sex predicted membership in the chronic trajectory relative to both the normative and escalating trajectories and also predicted desistance relative to escalation. In other words, males were somewhat overrepresented in the only group that exhibited a significant age-related decline, and this was reflected in the growth curve model. Notably, the chronic trajectory was nearly two-thirds male, consistent with Moffitt's conceptualization of a predominantly male life-course-persistent group.

Limitations

Our results should be interpreted in light of several limitations. First, participants were predominantly white. Although our sample was socioeconomically diverse and representative of the racial/ethnic makeup of the State of Michigan, future studies are needed to determine whether the present findings generalize to youth from more diverse racial/ethnic backgrounds. Next, participants were assessed on up to three occasions, meaning it was not possible to model non-linear fluctuations in RB. This may be particularly salient during adolescence, when engagement in RB is often observed to peak and subsequently decline (e.g., Doornwaard et al., 2012). Rather than follow a linearly increasing trajectory, for example, youth in the escalating group may have fluctuated in their engagement across adolescence. It is also unclear when, or if, youth in the escalating and chronic trajectories began to desist from RB. Results from prior studies that have followed participants into young adulthood have yielded mixed results as to when desistance occurs. In the Pittsburgh Youth Study, for instance, all trajectory groups exhibited near-zero levels of RB by the end of the study period (i.e., age 24) (Lacourse et al., 2008), whereas in a cohort from the Dunedin Multidisciplinary Health and Development Study, two of four trajectories (comprising nearly 1/3 of the sample) continued to engage in ASB at age 26 (Odgers et al., 2008). Subsequent studies with many assessment waves are needed to determine the typical timing of desistance from RB, as well as elucidate the factors predicting continued engagement into adulthood.

Next, we modeled RB based on reports from multiple informants at each wave. Although the use of multiple informants is likely to yield a fuller picture of RB than any single informant could provide (Achenbach, McConaughy, & Howell, 1987), youth self-reports were not available at Wave 1, and teacher reports were not available at Wave 3 (i.e., only caregiver reports were obtained at all waves). In addition to confirming measurement invariance, we conducted a series of sensitivity analyses using caregiver-reported RB as the outcome measure. As shown in Tables S3.7 and S3.8, as well as Figure S3.1, results were largely consistent with those from the main analyses. Results differed somewhat with respect to the frequency of engagement in RB during adolescence, as the unconditional growth curve indicated a slight *decrease* in RB with age, and a smaller percentage of participants followed an escalating trajectory than in the main analyses. These discrepancies likely reflect caregivers' incomplete awareness of their children's engagement in RB, as youth are both motivated to conceal these behaviors and increasingly able to do so as they get older (Burt, 2012). In support of this, meta-analytic work has found cross-informant correlations for a variety of psychiatric symptoms to be lower during adolescence than childhood (Achenbach et al., 1987). Nevertheless, the overall pattern of results was consistent, and peer delinquency remained a robust predictor of trajectory group membership.

Another limitation concerns the nested structure of the data. As noted in the Methods, we could not include the CLUSTER command in the conditional LGM or in any of the mixture models. We accounted for clustering in the conditional LGM by examining the data in "family-wide" format. As a verification, we also ran the conditional LGM with the data in "long" format, without controlling for clustering. As shown in Table S3.9, results were consistent with those from the main analyses. Notably, ADI predicted the intercept (in the expected direction), whereas it was marginally significant in the main analyses. For the mixture models, we ran the unconditional and conditional models with one twin per family randomly selected (see Table S3.10). The four-group solution was nearly identical to the solution from the full sample, with a modal trajectory of low/increasing RB, an escalating group, and two other groups distinguished by degree of desistance. Results were largely consistent in the conditional analyses as well, with elevated peer delinquency predicting membership in the escalating and chronic groups relative to the normative and desisting groups, respectively. Overall, we conclude that our results are not biased by the limitations of the software in accounting for clustering.

A methodological limitation specific to the mixture models was that the average posterior probability for the escalating trajectory was low. In mixture modeling, each participant is assigned to the group for which their posterior probability of membership is highest. Average posterior probabilities are thus a measure of classification accuracy, as higher values indicate greater certainty in class assignment. The average probability for the escalating group (.29 in the unconditional model and .59 in the conditional model) was likely related to the study's accelerated longitudinal design, with incomplete follow-up at the second and third waves of data collection. When the sample was restricted to those with complete data at all waves (N=377), the posterior probability for the escalating group was .73, and the probabilities for all other groups were > .80, above the recommended threshold of .70 (Nagin, 2005). The class solution was fully consistent with that observed in the full sample. In addition, average posterior probabilities for all groups but the escalating trajectory were > .70 in the full sample, suggesting that participants were assigned to groups with a reasonable degree of confidence, despite the planned missingness at later waves.

Lastly, although reports of neighborhood disadvantage, income, and peer delinquency were averaged across waves to represent each construct throughout the study period, they were considered as time-invariant predictors. Notably, the mean level of peer delinquency did not change significantly over time (b = -.001, p = .271), (although at Wave 2, youth reported significantly higher levels than caregivers did ($t(748) = 6.26$, $p < .001$), consistent with prior research on informant effects (Achenbach et al., 1987)). Although beyond the scope of this study, additional research is needed to elucidate the role of familial and neighborhood disadvantage and peer delinquency as time-varying predictors. In particular, studies should examine trajectories of peer delinquency and trajectories of RB in tandem to better understand the (likely reciprocal) relationship between the two.

Implications

Nevertheless, the present study has several important implications. First, the use of variable-centered and person-centered modeling approaches together yields a far more complete picture of development than either approach can provide by itself. Variable-centered methods are integral to modeling development across the full sample, which illuminates normative patterns and in so doing facilitates understanding of atypical development. In the present study, for example, the mean pattern in the growth curve indicated that a modest age-related increase in RB was normative, but that a sharp escalation was not. Variable-centered analyses were also useful for identifying predictors of elevated RB at baseline but were inconsistent in their ability to accurately model predictors of the slope. That is, some parameter estimates appeared to capture the actual association between predictor and slope (i.e., greater peer delinquency predicted a greater increase in RB), whereas others reflected regression to the mean artifacts (i.e., higher income predicted a greater increase in RB). Consideration of the person-centered results was necessary to determine which estimates from the growth curve were likely accurate and which were artifacts. Moreover, by allowing for the direct comparison of distinct trajectory groups, the person-centered analyses identified several factors predicting persistent and/or escalating engagement (i.e., low income, neighborhood deprivation, and peer delinquency), even though the sample likely did not comprise four qualitatively distinct trajectories (Walters & Ruscio, 2013). In short, both sets of analyses were needed to model the development of RB as a dimensional, yet highly heterogeneous, trait.

Second, the present study implicates peer delinquency as a critical risk factor for the development of RB. Across all analyses, peer delinquency was the most robust and consistent predictor, as it predicted both the intercept and slope of the growth curve (in the expected direction) and distinguished between five of the six group pairings from the mixture models. That greater peer delinquency predicted escalation relative to desistance suggests that its effects are particularly salient during adolescence, consistent with Moffitt's framework, which emphasizes the role of delinquent peer affiliation in the development of adolescent-limited ASB (Moffitt, 1993, 2003). Although the association between peer delinquency and RB is well-established (e.g., Ettekal & Ladd, 2015; Mann et al., 2015), no study to date has leveraged variable-centered and person-centered analyses to examine trajectories of peer delinquency and RB simultaneously. Increases in peer delinquency, for example, may predict increasing RB, evidenced by positive associations with the slope of RB and significant overlap with escalating RB trajectories. Given the unique strengths of each statistical approach, future studies that make use of both are needed to elucidate the role of peer delinquency over time, which may represent a promising target for prevention and intervention efforts addressing adolescent RB.

Overall, findings from both sets of analyses provide some additional support for Moffitt's framework, given the sample-wide and modal patterns of increasing RB reported here and the consistent evidence in support of peer influences. Moreover, the person-centered findings indicated that some participants followed a chronic, or life-course-persistent, trajectory, but that a larger percentage demonstrated onset in adolescence, consistent with Moffitt's original taxonomy (Moffitt, 1993). Despite these confirmations of the taxonomy's hypotheses, the present findings diverge from theory in two key ways. First, age-related increases in RB were far from universal, with a substantial portion of the sample demonstrating desistance. That is, of the 1,099 youth observed to exhibit childhood-onset RB, 694 (63%) recovered so completely that they engaged in less RB than even the normative group by the end of the study. Such findings suggest that recovery from early-onset ASB may be even more common than previously suggested (e.g., Moffitt et al., 2002). This raises questions about the nature of "life-course persistent" ASB. Second, the escalating group demonstrated such a dramatic increase that its engagement eventually surpassed that of the chronic group. This runs counter to Moffitt's conceptualization of adolescent-limited and life-course-persistent ASB as qualitatively distinct syndromes with marked differences in severity. The present findings instead point to differences between the two trajectories that are a matter of degree and suggest that some youth with late-onset ASB may become (at least) as impaired as those whose engagement begins in childhood. At the same time, some adolescent-limited youth likely conform to the pattern hypothesized by Moffitt, outgrowing their behavioral problems and successfully transitioning into adult roles (Moffitt, 1993). The likely distinctions *within* the classes that Moffitt proposed underscore the dimensional nature of RB (and ASB in general), as well as the need for both person-centered *and* variable-centered approaches to fully and accurately model trajectories across the early developmental period.

Tables

Table 3.1. *Descriptive statistics and correlations*

	1.	2.	3.	4.	5.	6.
1. ADI[a]	-					
2. Income[a]	**-.40**	-				
3. Peer Delinquency[b]	**.15**	**-.11**	-			
4. RB T1[c]	**.17**	**-.24**	**.32**	-		
5. RB T2[c]	**.14**	**-.13**	**.46**	**.31**	-	
6. RB T3[c]	.06	-.09	**.35**	**.30**	**.62**	-
Mean (SD)	56.60 (22.15)	6.91 (1.77)	6.32 (1.25)	2.54 (3.44)	2.74 (3.08)	2.91 (3.26)
Range	2-99	1-8	5-13	0-21	0-23	0-21
N	2031	2030	2021	2055	769	379

Note: Bold font indicates $p<.05$. [a]Mean scores were computed based on reports at all waves. [b]Mean scores were computed based on maternal and youth reports at all waves. [c]Multi-informant mean scores were computed based on maternal and teacher reports at Wave 1, maternal, teacher, and youth reports at Wave 2, and maternal and youth reports at Wave 3. Scores were multiplied by 24 to correspond to the maximum number of items at any wave.

Table 3.2. *Latent growth curve model fit statistics and parameter estimates*

Unconditional LGM model fit statistics					
Model	-2lnL	χ^2 (df)	AIC	BIC	SABIC
Linear growth	**7767.02**	**-**	**7779.02**	**7812.81**	**7793.74**
Means model	7834.63	67.61† (3)	7840.63	7857.52	7847.99
Conditional LGM parameter estimates					
Parameter	Estimate	S.E.	p-value		
Intercept					
Mean	**.644**	**.077**	**<.001**		
Variance	**.517**	**.042**	**<.001**		
ADI	.057	.030	.062		
Income	**-.162**	**.032**	**<.001**		
Peer Delinquency	**.170**	**.027**	**<.001**		
Sex	**.268**	**.052**	**<.001**		
Race	.059	.080	.460		
Slope					
Mean	**.030**	**.012**	**.010**		
Variance	**.006**	**.001**	**<.001**		
ADI	.002	.005	.713		
Income	**.012**	**.005**	**.016**		
Peer Delinquency	**.018**	**.005**	**<.001**		
Sex	**-.023**	**.008**	**.005**		
Race	-.002	.012	.878		
Covariances					
Intercept with slope	**-.046**	**.005**	**<.001**		
Intercept 1 with Intercept 2					
MZ	**.564**	**.040**	**<.001**		
DZ	**.365**	**.041**	**<.001**		
Slope 1 with Slope 2					
MZ	**.007**	**.001**	**<.001**		
DZ	**.005**	**.001**	**<.001**		
Intercept 1 with Slope 2					
MZ	**-.054**	**.005**	**<.001**		
DZ	**-.035**	**.005**	**<.001**		
Residual variance	**.290**	**.021**	**<.001**		

Note: Multi-informant composite mean RB scores are the outcome measure. Bold font indicates $p < .05$. †Significant change in chi-square at $p<.05$.

Table 3.3. *LCGA fit statistics and predictors of group membership*

Unconditional LCGA model fit statistics			
	AIC	BIC	Smallest group (% of sample)
# of groups			
1	7965.11	7982.00	100
2	7652.33	7686.11	38.9
3	7547.96	7598.63	6.0
4	7417.13	7484.69	7.2
5	7361.79	7446.24	4.4
6	7296.34	7397.67	2.3
Conditional LCGA parameter estimates from 4-group model			
Chronic v. normative			
	Estimate	S.E.	*p*-value
ADI	**.453**	**.135**	**.001**
Income	**-.634**	**.126**	**<.001**
Peer Delinquency	**2.072**	**.201**	**<.001**
Sex	**.868**	**.295**	**.003**
Race	.327	.314	.297
Escalating v. normative			
	Estimate	S.E.	*p*-value
ADI	**.460**	**.154**	**.003**
Income	-.159	.147	.279
Peer Delinquency	**1.762**	**.207**	**<.001**
Sex	-.277	.284	.330
Race	.271	.357	.448
Chronic v. desisting			
	Estimate	S.E.	*p*-value
ADI	.177	.140	.205
Income	**-.357**	**.119**	**.003**
Peer Delinquency	**1.212**	**.162**	**<.001**
Sex	.582	.301	.053
Race	.090	.311	.771

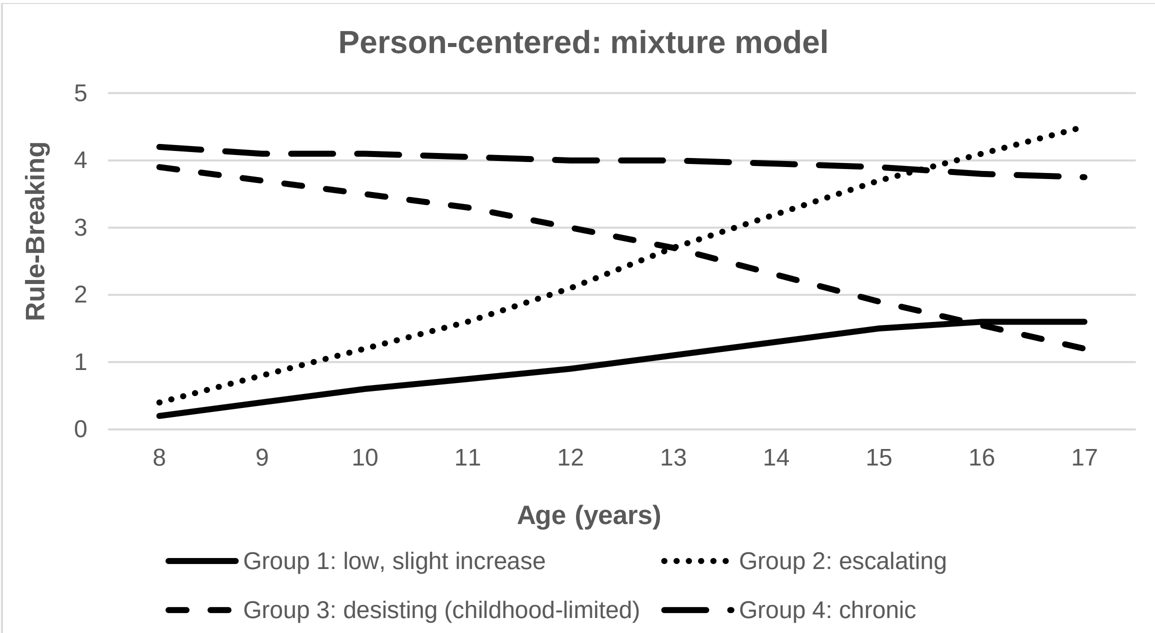
Note: Multi-informant composite mean RB scores are the outcome measure. Bold font indicates $p < .05$.

Table 3.4. *Characteristics of RB trajectory groups*

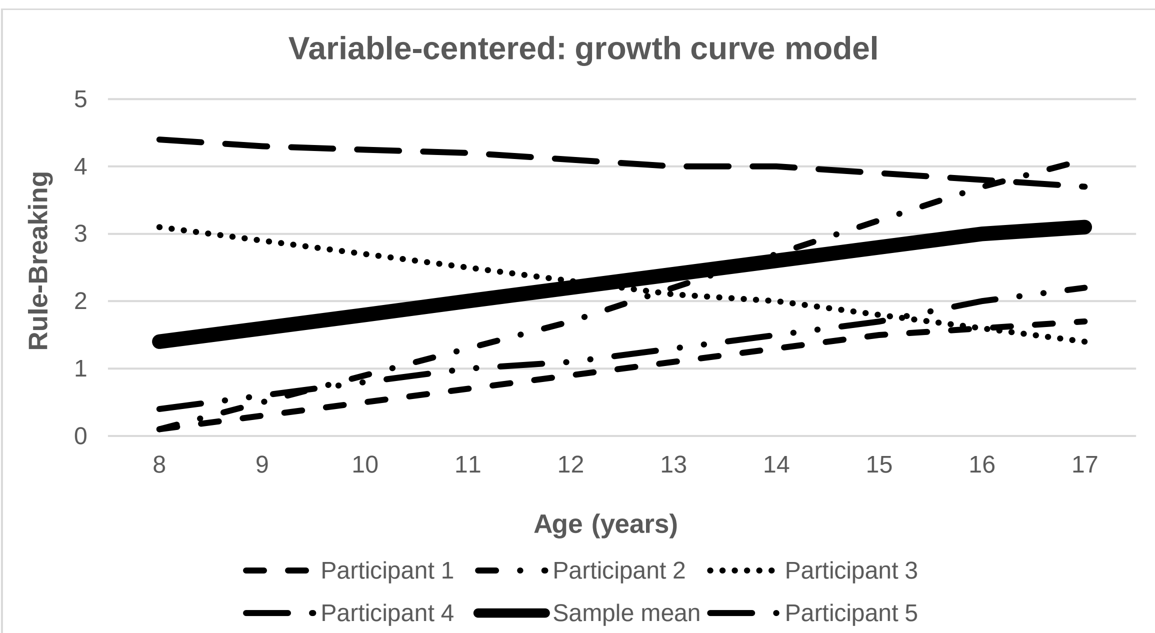
	Class 1: normative	Class 2: escalating	Class 3: desisting	Class 4: chronic
N	811	149	694	405
% male	42.5	47.7	54.8	64.4
		Means (SD) [Range]		
ADI	51.72 (21.39) [4-99]	64.03 (20.34) [21-99]	57.22 (22.30) [4-99]	62.42 (21.78) [2-99]
Income	7.23 (1.50) [1-8]	7.05 (1.47) [1.33-8]	6.89 (1.79) [1-8]	6.28 (2.15) [1-8]
Peer Delinquency	5.88 (1.04) [5-11]	6.57 (1.22) [5-13]	6.32 (1.17) [5-13]	7.11 (1.35) [5-11]
RB T1	.06 (.32) [0-3]	.21 (.52) [0-1.50]	2.95 (1.76) [0-15]	7.67 (3.72) [1.50-21]
RB T2	.69 (1.00) [0-7]	4.10 (2.93) [0-23]	1.51 (1.46) [0-13.50]	6.37 (3.29) [1-16.50]
RB T3	.77 (1.03) [0-4.50]	4.67 (3.39) [1.50-21]	1.74 (1.52) [0-7.50]	6.42 (3.72) [1.50-18]
RB T1 (log-transformed)	.04 (.18) [0-1.39]	.13 (.32) [0-.92]	1.30 (.38) [0-2.77]	2.07 (.43) [.92-3.09]
RB T2 (log-transformed)	.39 (.49) [0-2.08]	1.49 (.53) [0-3.18]	.78 (.54) [0-2.67]	1.91 (.43) [.69-2.86]
RB T3 (log-transformed)	.42 (.53) [0-1.70]	1.60 (.50) [.92-3.09]	.84 (.60) [0-2.14]	1.89 (.47) [.92-2.94]

Note: Multi-informant composite mean RB scores are the outcome measure.

Figures



a.



b.

Figure 3.1. *Person-centered (a) and variable-centered (b) approaches to modeling the development of RB in a hypothetical sample.* The person-centered model extracted four distinct trajectory groups, whereas the variable-centered model identified a mean-level increase across the entire sample.

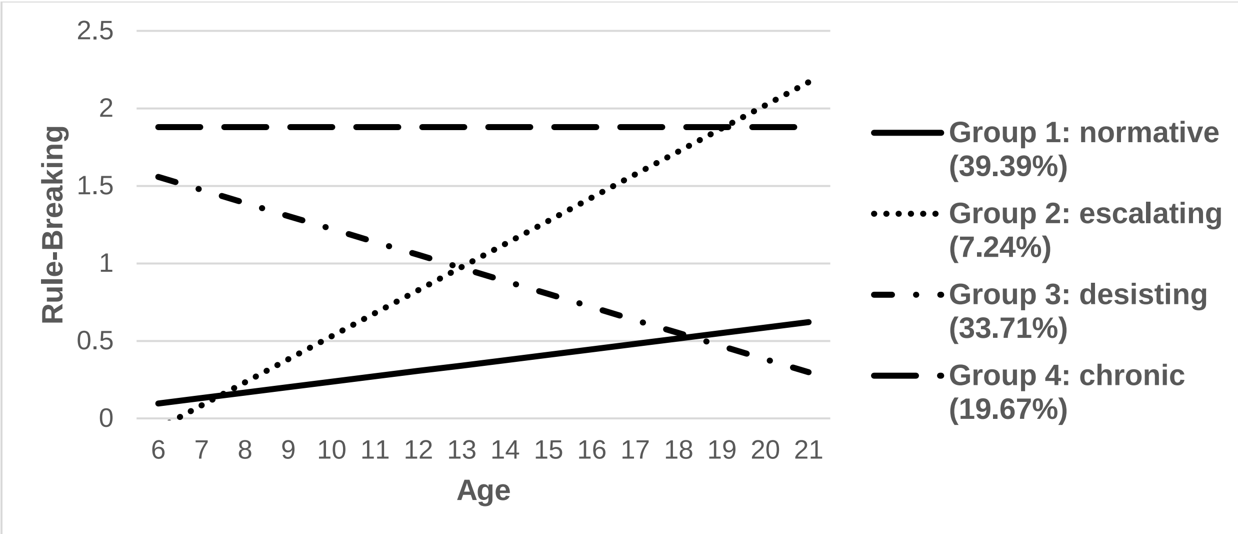


Figure 3.2. *Unconditional four-group model of RB trajectories based on multi-informant reports.*

GENERAL DISCUSSION

By leveraging multiple statistical approaches in a sample spanning nearly all of early development, the present studies substantially advanced our understanding of the origins of ASB. Study 1 found that more than one-third of the genetic and nonshared environmental contributions, respectively, to change in ASB were already present at the baseline assessment. Given the broad age range under study (i.e., ages three to 21 years), such findings indicate a remarkable degree of continuity over time in both genetic and environmental influences. Furthermore, the present findings align with prior research on youth psychopathology indicating that the most potent environmental influences are those that make children in the same family less alike (Plomin & Daniels, 1987). Study 1 clearly implicates (some of) these influences as enduring and systematic, rather than merely transient, and proposes several possible nonshared environmental factors that may systematically shape the development of ASB, including differential parenting, delinquent peer affiliation, and aspects of the family environment that are objectively shared but impact siblings in distinct ways (e.g., parental divorce; Goldsmith, 1993).

Although the analyses were not genetically informed, Studies 2 and 3 illuminated several putatively environmental contributions to trajectories of AGG and RB through their use of both variable-centered and person-centered methods. In Study 2, parent-child conflict predicted not only baseline engagement in AGG, but also change over time, such that youth experiencing less conflict with their mothers were far more likely to exhibit desistance. Put differently, low levels of conflict appeared to interrupt trajectories of elevated AGG, consistent with prior work identifying parent-child conflict as a key psychosocial risk factor for externalizing (Burt, Krueger, McGue, & Iacono, 2003). The prospect of differential parenting as a systematic, nonshared environmental influence on ASB trajectories is also bolstered by studies leveraging quasi-experimental designs, which have found twin differences in parent-child conflict (Burt, McGue, Iacono, & Krueger, 2006) and harsh parenting (Burt et al., 2021) to predict twin differences in

ASB. Such findings support an association between parenting and ASB that is at least partially environmental in origin and may well be causal to some extent. Study 2 builds on this prior work by implicating the *absence* of conflict as a protective factor for youth who initially demonstrate high levels of AGG, although additional research leveraging genetically informative designs is needed to clarify whether the association between low conflict and desistance is also environmentally mediated.

In turn, Study 3 found both neighborhood disadvantage and affiliation with delinquent peers to predict RB trajectories, consistent with prior research on the effects of the broader neighborhood (e.g., Lacourse et al., 2008) and peer (e.g., Mann et al., 2015) contexts. Moreover, the effects of neighborhood disadvantage were observed when controlling for household income, indicating a unique role of contextual disadvantage in the development of RB (consistent with findings from cross-sectional studies that disambiguated neighborhood and familial disadvantage (Carroll et al., 2023; Kupersmidt et al., 1995)). Although objectively shared by youth residing in the same home, the neighborhood context may impact each child in idiosyncratic ways. Neighborhood disadvantage is also a multifaceted construct (Singh, 2003), encompassing risk factors such as pervasive poverty, household overcrowding, and community violence, which may be shared by siblings to differing degrees. Differences between siblings in exposure to community violence, for instance, might predict differences in their engagement in RB. The association between peer delinquency and RB, on the other hand, may be more complex, given that youth are able to select into peer groups in a way that they do not typically select into neighborhoods. Studies leveraging cross-lagged designs have sought to clarify whether the association is more indicative of selection (i.e., youth who engage in ASB seek out delinquent peers) or socialization (i.e., peer delinquency confers risk for ASB). Although some work has pointed solely to a selection effect (Burt, McGue, & Iacono, 2009), most studies have found evidence of both selection and socialization (Dugré, Giguére, & Potvin, 2024; Kendler, Jacobson, Myers, & Eaves, 2008; Samek et al., 2016; Schwartz, Solomon, & Valgardson, 2019). Other work has indicated that the respective roles of selection and socialization are age-dependent, with greater evidence of selection effects during adolescence relative to childhood (Kendler et al., 2008) as youth are increasingly able to select into environments that are consistent with their predispositions (Beam & Turkheimer, 2013). Additional longitudinal studies, including those incorporating genetically informative designs, are needed to elucidate the temporal associations between peer delinquency and RB (and ASB in general) throughout the early developmental period.

Some of the predictors discussed above may not be specific to either AGG or RB but rather confer risk for trajectories of ASB in general. Indeed, Burt and colleagues found parent-child conflict to underlie comorbidity for multiple forms of externalizing (Burt et al., 2003), and Dugré et al. (2024) reported bidirectional associations between peer delinquency and AGG and RB, respectively. Furthermore, AGG and RB often co-occur, with correlations of moderate-to-large magnitude typically reported between the two (*r*s .4-.6; Burt, 2013). Such findings raise the possibility that the high-risk trajectories identified via mixture modeling in Studies 2 and 3 largely comprise the same participants. Cross-tabulation results, shown in Table 4.1, help to clarify the extent of the overlap between trajectories of AGG and trajectories of RB in the present sample. As expected, participants evidenced some overlap in their group memberships; most participants who were assigned to the normative trajectory for one dimension of ASB, for instance, were also assigned to the normative trajectory for the other. In addition, nearly two-thirds of those following a persistent AGG trajectory also exhibited persistent RB. However, the reverse was not true, as most participants who exhibited persistent RB followed either normative or desisting trajectories of AGG. In addition, most youth who followed desisting trajectories of one form of ASB did *not* do so for the other. Lastly, the majority of those following a trajectory of escalating RB, for which there was not an analogous trajectory of AGG, fell into the normative AGG group.

These findings collectively indicate that associations between AGG and RB trajectories are far from unity, consistent with results from the few extant studies examining overlapping trajectories (e.g., Mata & van Dulmen, 2012; Maughan et al., 2000). Far more work is needed to understand how AGG and RB develop in tandem as related yet distinct traits. Studies should leverage parallel process growth curve models to examine the development of AGG and RB, respectively, across entire samples, as well as parallel process mixture models to capture patterns of growth for distinct subgroups. As discussed above, it is also unclear which risk/protective factors are specific to each dimension of ASB and which confer, or reduce, risk for ASB in general. Of the 1078 youth following normative trajectories of AGG, for instance, why do 445 exhibit non-normative trajectories of RB? Moreover, how do these youth differ from those following persistent trajectories of both AGG and RB? Future research could answer these questions by examining predictors of joint trajectories via mixture modeling or via analyses that control for one dimension of ASB while examining engagement in the other.

There are several other limitations to the present studies. First, all studies focused on changes in the frequency, rather than severity, of ASB, and some of the behaviors assessed were of relatively low severity (e.g., hot temper, breaks rules). Among those youth engaging in frequent and persistent ASB, severity typically increases with age (e.g., Copeland, Miller-Johnson, Keeler, Angold, & Costello, 2007). Additional research is needed to elucidate when and how escalations in severity occur, as well as which specific behaviors comprise the highest-risk trajectories. In addition, because comorbidity is associated with greater severity (e.g., Kessler, Chiu, Demler, & Walters, 2005), future studies should examine patterns of comorbid AGG and RB in conjunction with age-related changes in the specific types of offenses committed.

Next, it is unclear to what extent the effects of the risk and protective factors under study differed across sex. The present studies found males to exhibit higher levels of AGG, RB, and ASB in general at baseline, with fewer differences across sex observed in rates of change, consistent with prior work (e.g., Miller, Malone, & Dodge, 2010). However, we did not examine potential interactions between sex and familial disadvantage, neighborhood disadvantage, parent-child conflict, or peer delinquency. Prior cross-sectional work has found associations between disadvantage and externalizing to persist regardless of whether sex was included as a covariate (Carroll et al., 2023), but it is unclear whether the same would be true longitudinally. There is some evidence that associations between the parent-child relationship and trajectories of ASB are similar for males and females (Janssen, Eichelsheim, Dekovic, & Bruinsma, 2017; Snyder, Schrepferman, Bullard, McEachern, & Patterson, 2012), but these studies followed participants across relatively brief windows of development. Lastly, regarding peer delinquency, some studies have found associations between peer influences and ASB to be stronger for males (e.g., Piquero, Gover, MacDonald, & Piquero, 2005), whereas others have found females to be more susceptible to the influences of delinquent peers (e.g., Haynie, Doogan, & Soller, 2014). Haynie and colleagues also reported sex differences in the respective roles of selection and socialization, with stronger evidence of peer selection effects for girls. In short, much remains unknown about sex differences in the developmental origins of ASB. Future longitudinal studies should leverage variable-centered and person-centered methods, as well as quasi-experimental designs, to examine interactions between sex and other predictors of AGG and RB trajectories.

Lastly, none of the present studies examined the role of genotype-by-environment interaction (GxE) in the development of ASB. GxE refers to the differential impact of environmental risk or protective factors as a function of genotype (and vice versa, i.e., environmental factors impact the expression of genetic predispositions) (Plomin, DeFries, & Loehlin, 1977). Extant cross-sectional work has implicated GxE as an important contributor to youth ASB, particularly RB. For example, Burt and colleagues found shared environmental influences on RB to be amplified, and genetic contributions suppressed, for youth residing in disadvantaged neighborhoods (Burt, Klump, Gorman-Smith, & Neiderhiser, 2016). This pattern of findings, which has been consistent across studies (Cleveland, 2003; Tuvblad, Grann, &

Lichtenstein, 2006), is indicative of a bioecological GxE, with deleterious environments exerting such a strong effect on development that poor mental health outcomes are observed even among youth who are not genetically predisposed (Bronfenbrenner & Ceci, 1994).

Despite these robust and consistent findings linking GxE to the *occurrence* of ASB, the role of GxE in the *development* of ASB is unknown, as models to assess GxE longitudinally do not yet exist (though they are in progress). Given age-related changes in both the frequency (Moffitt, 1993) and the etiology (Rhee & Waldman, 2002) of ASB, there are a variety of ways in which longitudinal findings may differ from cross-sectional ones. First, the manifestation of bioecological GxE may shift from one involving amplification of shared environmental influences to an increase in the effects of the *non*shared environment. This developmental shift would be consistent with the age-related increase in nonshared environmental influences, and co-occurring decrease in shared environmental effects, on ASB reported in prior meta-analytic work (Rhee & Waldman, 2002), which is hypothesized to reflect adolescents' increasing ability to select their own environments outside the family. As youth progress into emerging adulthood, they may thus be less susceptible to familial influences but more affected by experiences in the neighborhood context that are unique to each person in a given family. Alternatively, GxE may follow an entirely different pattern. The diathesis-stress model of GxE posits that genetic risk would be activated in deleterious environments (Ingram & Luxton, 2005), meaning that genetic contributions to ASB would be expected to be *amplified* in disadvantaged neighborhoods. This pattern of GxE stands in direct contrast to that proposed by the bioecological model and suggests another way in which genetic and environmental influences may interact to shape youth behavioral trajectories. A shift from bioecological GxE during childhood to diathesis-stress GxE in adolescence would also help to explain age-related increases in the heritability of RB (Burt, 2015). In short, longitudinal studies are needed to understand GxE as a developmental process unfolding over time, rather than a static phenomenon. Future work should leverage both variable-centered and person-centered approaches to elucidate the effects of neighborhood disadvantage on the etiology of the intercept and slope of ASB, as well as trajectory group membership.

Despite these limitations, the present studies yield several important conclusions regarding the development of ASB. First, behavioral trajectories are shaped by genetic and nonshared environmental influences that are partially continuous across almost all of early development. That is, some of the same etiological factors that contribute to ASB in emerging adulthood are present as early as the preschool years. Some of these factors are person-specific environmental, as the nonshared environment comprised not only transient, unsystematic effects but also enduring influences that shaped the magnitude of age-related change. Regarding specific predictors of ASB trajectories, the present studies implicated individual (i.e., biological sex), familial (i.e., parent-child conflict, household income), social network (i.e., peer delinquency), and neighborhood (i.e., concentrated deprivation) characteristics, consistent with theoretical work viewing development to unfold via interactions between the individual and numerous proximal and distal environmental contexts (Bronfenbrenner, 1988).

Second, the nuanced ways in which Studies 2 and 3 were able to illuminate the predictors of desistance and escalation, respectively, underscore the need to leverage both variable-centered and person-centered statistical methods when modeling trajectories of psychopathology. Across all three studies, variable-centered methods were severely limited in their ability to model predictors of the slope, with neighborhood disadvantage and parent-child conflict found to predict *declining* ASB and household income found to predict an age-related *increase*. It was only by incorporating person-centered models that we were able to accurately conceptualize the respective roles of socioeconomic disadvantage and parent-child conflict in predicting rates of change. At the same time, taxometric work has consistently found ASB to exist as a continuous trait (Walters, 2011; Walters & Ruscio, 2013), like other forms of psychopathology (Krueger & Piasecki, 2002). Indeed, participants in the present studies were not assigned to trajectory groups with absolute certainty, in part because the groups likely represent approximations of participants' trajectories rather than qualitatively distinct entities. By leveraging both approaches together, the present studies were able to identify risk/protective factors associated with specific ASB trajectories, while simultaneously modeling the development of ASB in a way that was consistent with its underlying distribution. Future studies that make use of both statistical approaches, in conjunction with genetically informative designs that examine GxE, will further elucidate the origins of ASB and other forms of developmental psychopathology.

Tables

Table 4.1. *Cross-tabulation results for AGG and RB trajectory group membership*

		AGG			**Total *N***
		Normative	Desisting	Persisting	
RB	Normative	633	168	6	807
	Escalating	100	29	20	149
	Desisting	287	322	80	689
	Persisting	58	161	184	403
	Total *N*	1078	680	290	2048

BIBLIOGRAPHY

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological bulletin*, *101*(2), 213.

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms & profiles. *Burlington, VT: University of Vermont Research Centre for Children, Youth and Families*.

Akaike, H. (1987). Factor analysis and AIC. In *Selected Papers of Hirotugu Akaike* (pp. 371-386). Springer, New York, NY.

Alink, L. R., & Egeland, B. (2013). The roles of antisocial history and emerging adulthood developmental adaptation in predicting adult antisocial behavior. *Aggressive behavior*, *39*(2), 131-140.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Arım, R. G., Dahinten, V. S., Marshall, S. K., & Shapka, J. D. (2011). An examination of the reciprocal relationships between adolescents' aggressive behaviors and their perceptions of parental nurturance. *Journal of Youth and Adolescence*, *40*, 207-220.

Barker, E. D., Séguin, J. R., White, H. R., Bates, M. E., Lacourse, E., Carbonneau, R., & Tremblay, R. E. (2007). Developmental trajectories of male physical violence and theft: Relations to neurocognitive performance. *Archives of General Psychiatry*, *64*(5), 592–599.

Bartels, M., Van den Oord, E. J. C. G., Hudziak, J. J., Rietveld, M. J. H., Van Beijsterveldt, C. E. M., & Boomsma, D. I. (2004). Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. *Developmental psychology*, *40*(5), 852.

Bauer, D. J., & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: potential problems and promising opportunities. *Psychological methods*, *9*(1), 3.

Beam, C. R., & Turkheimer, E. (2013). Phenotype–environment correlations in longitudinal twin models. *Development and psychopathology*, *25*(1), 7-16.

Beauchaine, T. P., Webster-Stratton, C., & Reid, M. J. (2005). Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: a latent growth curve analysis. *Journal of consulting and clinical psychology*, *73*(3), 371.

Becht, A. I., Prinzie, P., Deković, M., van den Akker, A. L., & Shiner, R. L. (2016). Child personality facets and overreactive parenting as predictors of aggression and rule-breaking trajectories from childhood to adolescence. *Development and Psychopathology*, *28*(2), 399–413.

Becker, A., Woerner, W., Hasselhorn, M., Banaschewski, T., & Rothenberger, A. (2004).

Validation of the parent and teacher SDQ in a clinical sample. *European child & adolescent psychiatry*, *13*(2), ii11-ii16.

Benson, M. J., & Buehler, C. (2012). Family process and peer deviance influences on adolescent aggression: Longitudinal effects across early and middle adolescence. *Child development*, *83*(4), 1213-1228.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, *88*(3), 588.

Berkout, O. V., Young, J. N., & Gross, A. M. (2011). Mean girls and bad boys: Recent research on gender differences in conduct disorder. *Aggression and Violent Behavior*, *16*(6), 503-511.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological methods & research*, *17*(3), 303-316.

Bongers, I. L., Koot, H. M., Van Der Ende, J., & Verhulst, F. C. (2004). Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child development*, *75*(5), 1523-1537.

Brame, B., Nagin, D. S., & Tremblay, R. E. (2001). Developmental trajectories of physical aggression from school entry to late adolescence. *Journal of child psychology and psychiatry*, *42*(4), 503-512.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., Fergusson, D., Horwood, J. L., Loeber, R., Laird, R., Lynam, D. R., Moffitt, T. E., Pettit, G. S., & Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, *39*(2), 222–245.

Bronfenbrenner, U. (1988). Interacting systems in human development. Research paradigms: Present and future. Persons in context: Developmental processes, 2, 25-49.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological review*, *101*(4), 568.

Burt, S. A. (2009). Rethinking environmental contributions to child and adolescent psychopathology: a meta-analysis of shared environmental influences. *Psychological bulletin*, *135*(4), 608.

Burt, S. A. (2012). How do we optimally conceptualize the heterogeneity within antisocial behavior? An argument for aggressive versus non-aggressive behavioral dimensions. *Clinical psychology review*, *32*(4), 263-279.

Burt, S. A. (2013). Do etiological influences on aggression overlap with those on rule breaking? A meta-analysis. *Psychological Medicine*, *43*(9), 1801-1812.

Burt, S. A. (2015). Evidence that the gene–environment interactions underlying youth conduct problems vary across development. *Child Development Perspectives*, *9*(4), 217-221.

Burt, S. A. (2022). The genetic, environmental, and cultural forces influencing youth antisocial behavior are tightly intertwined. *Annual review of clinical psychology*, *18*, 155-178.

Burt, S. A., Clark, D. A., Gershoff, E. T., Klump, K. L., & Hyde, L. W. (2021). Twin Differences in Harsh Parenting Predict Youth's Antisocial Behavior. *Psychological science*, *32*(3), 395-409.

Burt, S. A., Brent Donnellan, M., Slawinski, B. L., & Klump, K. L. (2016). The phenomenology of non-aggressive antisocial behavior during childhood. *Journal of abnormal child psychology*, *44*, 651-661.

Burt, S. A., Klahr, A. M., & Klump, K. L. (2015). Do non-shared environmental influences persist over time? An examination of days and minutes. *Behavior genetics*, *45*(1), 24-34.

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

Burt, S. A., & Klump, K. L. (2019). The Michigan State University Twin Registry (MSUTR): 15 years of twin and family research. *Twin research and human genetics: the official journal of the International Society for Twin Studies*, *22*(6), 741.

Burt, S. A., Klump, K. L., Gorman-Smith, D., & Neiderhiser, J. M. (2016). Neighborhood disadvantage alters the origins of children's nonaggressive conduct problems. *Clinical psychological science*, *4*(3), 511-526.

Burt, S. A., Krueger, R. F., McGue, M., & Iacono, W. (2003). Parent-child conflict and the comorbidity among childhood externalizing disorders. *Archives of general psychiatry*, *60*(5), 505-513.

Burt, S. A., McGue, M., Carter, L. A., & Iacono, W. G. (2007). The different origins of stability and change in antisocial personality disorder symptoms. *Psychological Medicine*, *37*(1), 27.

Burt, S. A., McGue, M., & Iacono, W. G. (2009). Nonshared environmental mediation of the association between deviant peer affiliation and adolescent externalizing behaviors over time: results from a cross-lagged monozygotic twin differences design. *Developmental psychology*, *45*(6), 1752.

Burt, S. A., McGue, M., Iacono, W. G., & Krueger, R. F. (2006). Differential parent-child relationships and adolescent externalizing symptoms: cross-lagged analyses within a monozygotic twin differences design. *Developmental psychology*, *42*(6), 1289.

Burt, S. A., Slawinski, B. L., Carsten, E. E., Harden, K. P., Hyde, L. W., & Klump, K. L. (2019). How should we understand the absence of sex differences in the genetic and environmental origins of antisocial behavior?. *Psychological medicine*, *49*(10), 1600.

Carroll, S. L., Mikhail, M. E., & Burt, S. A. (2023). The development of youth antisocial behavior across time and context: A systematic review and integration of person-centered and variable-centered research. *Clinical psychology review*, 102253.

Carroll, S. L., Shewark, E. A., Mikhail, M. E., Thaler, D. J., Pearson, A. L., Klump, K. L., & Burt, S. A. (2023). Identifying the 'active ingredients' of socioeconomic disadvantage for youth outcomes in middle childhood. *Development and Psychopathology*, 1-9.

Cleveland, H. H. (2003). Disadvantaged neighborhoods and adolescent aggression: Behavioral genetic evidence of contextual effects. *Journal of Research on Adolescence*, *13*(2), 211-238.

Conger, R. D., Ge, X., Elder Jr, G. H., Lorenz, F. O., & Simons, R. L. (1994). Economic stress, coercive family process, and developmental problems of adolescents. *Child development*, *65*(2), 541-561.

Copeland, W. E., Miller-Johnson, S., Keeler, G., Angold, A., & Costello, E. J. (2007). Childhood psychiatric disorders and young adult crime: A prospective, population-based study. *American Journal of psychiatry*, *164*(11), 1668-1675.

Côté, S., Vaillancourt, T., LeBlanc, J. C., Nagin, D. S., & Tremblay, R. E. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A nation wide longitudinal study of Canadian children. *Journal of abnormal child psychology*, *34*, 68-82.

DeGarmo, D. S., & Forgatch, M. S. (2005). Early development of delinquency within divorced families: Evaluating a randomized preventive intervention trial. *Developmental science*, *8*(3), 229-239.

DeGarmo, D. S., Patterson, G. R., & Forgatch, M. S. (2004). How do outcomes in a specified parent training intervention maintain or wane over time?. *Prevention Science*, *5*, 73-89.

de Haan, A. D., Prinzie, P., & Deković, M. (2012). Change and reciprocity in adolescent aggressive and rule-breaking behaviors and parental support and dysfunctional discipline. *Development and Psychopathology*, *24*(1), 301–315.

Diamantopoulou, S., Verhulst, F. C., & van der Ende, J. (2011). Gender differences in the development and adult outcome of co-occurring depression and delinquency in adolescence. *Journal of Abnormal Psychology*, *120*(3), 644–655.

Doornwaard, S. M., Branje, S., Meeus, W. H. J., & ter Bogt, T. F. M. (2012). Development of adolescents' peer crowd identification in relation to changes in problem behaviors. *Developmental Psychology*, *48*(5), 1366–1380.

Dong, L. (2016). The role of perinatal depression in the development of co-occurring aggression and anxiety/depression problems [PhD Thesis, Emory University]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering*.

Dugré, J. R., Giguère, C. É., & Potvin, S. (2024). The company you keep: the neglected role of affiliating with delinquent friends in the development of the cannabis-violence link. *Addictive behaviors*, *151*, 107939.

Eamon, M. K. (2002). Poverty, parenting, peer, and neighborhood influences on young adolescent antisocial behavior. *Journal of Social Service Research*, *28*(1), 1-23.

Ehrenreich, S. E., Beron, K. J., Brinkley, D. Y., & Underwood, M. K. (2014). Family predictors of continuity and change in social and physical aggression from ages 9 to 18. *Aggressive Behavior*, *40*(5), 421–439.

Eley, T. C., Lichtenstein, P., & Moffitt, T. E. (2003). A longitudinal behavioral genetic analysis of the etiology of aggressive and nonaggressive antisocial behavior. *Development and psychopathology*, *15*(2), 383-402.

Elkins, I. J., McGue, M., & Iacono, W. G. (1997). Genetic and environmental influences on parent–son relationships: Evidence for increasing genetic influence during adolescence. *Developmental psychology*, *33*(2), 351.

Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology 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*(3), 430-457.

El-Sheikh, M., & Elmore–Staton, L. O. R. I. (2004). The link between marital conflict and child adjustment: Parent–child conflict and perceived attachments as mediators, potentiators, and mitigators of risk. *Development and Psychopathology*, *16*(3), 631-648.

Ettekal, I., & Ladd, G. W. (2015). Developmental pathways from childhood aggression–disruptiveness, chronic peer rejection, and deviant friendships to early-adolescent rule breaking. *Child development*, *86*(2), 614-631.

Ferdinand, R. F., Hoogerheide, K. N., Van Der Ende, J., Visser, J. H., Koot, H. M., Kasius, M. C., & Verhulst, F. C. (2003). The role of the clinician: three-year predictive value of parents', teachers', and clinicians' judgment of childhood psychopathology. *Journal of Child Psychology and Psychiatry*, *44*(6), 867-876.

Fontaine, N. M., Lacourse, E., Vitaro, F., & Tremblay, R. E. (2014). Links between trajectories of self-reported violent and nonviolent offending and official offending during adolescence and adulthood. *Criminal behaviour and mental health*, *24*(4), 277-290.

Ganiban, J. M., Saudino, K. J., Ulbricht, J., Neiderhiser, J. M., & Reiss, D. (2008). Stability and change in temperament during adolescence. *Journal of personality and social psychology*, *95*(1), 222.

Gard, A. M., Waller, R., Shaw, D. S., Forbes, E. E., Hariri, A. R., & Hyde, L. W. (2017). The long reach of early adversity: Parenting, stress, and neural pathways to antisocial behavior in adulthood. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *2*(7), 582-590.

Gardner, T. W. (2006). *Developmental trajectories of antisocial behavior: Prediction of academic failure, peer rejection, and depression in adolescence*. University of Oregon.

Gardner, T. W., Dishion, T. J., & Connell, A. M. (2008). Adolescent self-regulation as resilience: Resistance to antisocial behavior within the deviant peer context. *Journal of abnormal child psychology*, *36*(2), 273-284.

Goldsmith, H. H. (1993). Nature-nurture issues in the behavioral genetic context: Overcoming barriers to communication. *Nature, nurture, and psychology*, 325-339.

Goodman, R. (2001). Strengths and Difficulties Questionnaire (SDQ). Youthinmind, London. Retreived 20 August 2007.

Grimm, K. J., Ram, N., Estabrook, R. (2017). Growth modeling: Structural equation and multilevel modeling approaches. The Guilford Press.

Hagen, K. A., Ogden, T., & Bjørnebekk, G. (2011). Treatment outcomes and mediators of parent management training: A one-year follow-up of children with conduct problems. *Journal of Clinical Child & Adolescent Psychology*, *40*(2), 165-178.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Factor analysis. Multivariate data analysis. *NJ: Prentice-Hall*, *3*, 98-99.

Haynie, D. L., Doogan, N. J., & Soller, B. (2014). Gender, friendship networks, and delinquency: A dynamic network approach. *Criminology*, *52*(4), 688-722.

He, J. P., Burstein, M., Schmitz, A., & Merikangas, K. R. (2013). The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in US adolescents. *Journal of abnormal child psychology*, *41*(4), 583-595.

Hirschi, T., & Gottfredson, M. R. (1983). *Age and the Explanation of Crime*. American Journal of Sociology, 89, 552–584.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Evaluating model fit: a synthesis of the structural equation modelling literature. In *7th European Conference on research methodology for business and management studies* (Vol. 2008, pp. 195-200).

Hopwood, C. J., Donnellan, M. B., Blonigen, D. M., Krueger, R. F., McGue, M., Iacono, W. G., & Burt, S. A. (2011). Genetic and environmental influences on personality trait stability and growth during the transition to adulthood: a three-wave longitudinal study. *Journal of personality and social psychology*, *100*(3), 545.

Hox, J., & Stoel, R. D. (2014). Multilevel and SEM approaches to growth curve modeling. *Wiley StatsRef: Statistics Reference Online*.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Ingoldsby, E. M. (2003). Neighborhood contextual factors and early-starting antisocial behavior [PhD Thesis]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering*.

Ingram, R. E., & Luxton, D. D. (2005). Vulnerability-stress models. *Development of psychopathology: A vulnerability-stress perspective*, *46*(2), 32-46.

Isen, J., Tuvblad, C., Younan, D., Ericson, M., Raine, A., & Baker, L. A. (2022). Developmental trajectories of delinquent and aggressive behavior: Evidence for differential heritability. *Child Psychiatry & Human Development*, *53*(2), 199-211.

Jacobson, K. C., Prescott, C. A., & Kendler, K. S. (2002). Sex differences in the genetic and environmental influences on the development of antisocial behavior. *Development and Psychopathology*, *14*(2), 395-416.

Janssen, H. J., Eichelsheim, V. I., Deković, M., & Bruinsma, G. J. (2017). Sex differences in longitudinal pathways from parenting to delinquency. *European Journal on Criminal Policy and Research*, *23*, 503-521.

Karriker-Jaffe, K. J., Foshee, V. A., Ennett, S. T., & Suchindran, C. (2008). The development of aggression during adolescence: Sex differences in trajectories of physical and social aggression among youth in rural areas. *Journal of Abnormal Child Psychology*, *36*(8), 1227–1236.

Kendler, K. S., Jacobson, K., Myers, J. M., & Eaves, L. J. (2008). A genetically informative developmental study of the relationship between conduct disorder and peer deviance in males. *Psychological medicine*, *38*(7), 1001-1011.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry*, *62*(6), 617-627.

Kind, A. J., & Buckingham, W. R. (2018). Making neighborhood-disadvantage metrics accessible—the neighborhood atlas. *The New England journal of medicine*, *378*(26), 2456.

Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: Reducing child conduct problems in real-world settings. Behaviour Research and Therapy, 51(3), 113– 121.

Klahr, A. M., Klump, K. L., & Burt, S. A. (2014). The etiology of the association between child antisocial behavior and maternal negativity varies across aggressive and non-aggressive rule-breaking forms of antisocial behavior. *Journal of abnormal child psychology*, *42*, 1299-1311.

Krueger, R. F., & Piasecki, T. M. (2002). Toward a dimensional and psychometrically-informed approach to conceptualizing psychopathology. *Behaviour research and therapy*, *40*(5), 485-499.

Kupersmidt, J. B., Griesler, P. C., DeRosier, M. E., Patterson, C. J., & Davis, P. W. (1995). Childhood aggression and peer relations in the context of family and neighborhood factors. *Child development*, *66*(2), 360-375.

Lacourse, E., Boivin, M., Brendgen, M., Petitclerc, A., Girard, A., Vitaro, F., ... & Tremblay, R. E. (2014). A longitudinal twin study of physical aggression during early childhood: evidence for a developmentally dynamic genome. *Psychological medicine*, *44*(12), 2617.

Lacourse, E., Dupéré, V., & Loeber, R. (2008). Developmental trajectories of violence and theft. In R. Loeber, D. P. Farrington, M. Stouthamer-Loeber, & H. R. White (Eds.), *Violence and serious theft: Development and prediction from childhood to adulthood BT - Violence and serious theft: Development and prediction from childhood to adulthood* (pp. 231–268, Chapter xxiii, 404 Pages). Routledge/Taylor & Francis Group, New York, NY.

Lahey, B. B., Loeber, R., Burke, J. D., & Applegate, B. (2005). Predicting future antisocial personality disorder in males from a clinical assessment in childhood. *Journal of consulting and clinical psychology*, *73*(3), 389.

Lee, Y., Liu, X., & Watson, M. W. (2016). The timing effect of bullying in childhood and adolescence on developmental trajectories of externalizing behaviors. *Journal of interpersonal violence*, *31*(17), 2775-2800.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. *Psychological bulletin*, *126*(2), 309.

Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (Vol. 793). John Wiley & Sons.

MacCallum, R. C. (2009). Factor Analysis. In R. E. Millsap, & A. Maydeu-Olivares (Eds.), *The SAGE Handbook of Quantitative Methods in Psychology* (pp. 123-147). London: SAGE Publications Ltd.

Maldonado-Molina, M. M., Reingle, J. M., Tobler, A. L., Jennings, W. G., & Komro, K. A. (2010). Trajectories of physical aggression among Hispanic urban adolescents and young adults: An application of latent trajectory modeling from ages 12 to 18. *American Journal of Criminal Justice*, *35*, 121-133.

Mann, F. D., Kretsch, N., Tackett, J. L., Harden, K. P., & Tucker-Drob, E. M. (2015). Person× environment interactions on adolescent delinquency: Sensation seeking, peer deviance and parental monitoring. *Personality and individual differences*, *76*, 129-134.

Marsh, H. W., & Hau, K. T. (2002). Multilevel modeling of longitudinal growth and change: Substantive effects or regression toward the mean artifacts?. *Multivariate Behavioral Research*, *37*(2), 245-282.

Martino, S. C., Ellickson, P. L., Klein, D. J., McCaffrey, D., & Edelen, M. O. (2008). Multiple trajectories of physical aggression among adolescent boys and girls. *Aggressive behavior*, *34*(1), 61-75.

Mata, A. D. (2013). Structured activity involvement and behavior problems: Investigation of selection and transactional effects with the use of variable and person-centered approaches [PhD Thesis]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering*.

Mata, A. D., & Van Dulmen, M. H. (2012). Adult-onset antisocial behavior trajectories: Associations with adolescent family processes and emerging adulthood functioning. *Journal of interpersonal violence*, *27*(1), 177-193.

Maughan, B., Pickles, A., Rowe, R., Costello, E. J., & Angold, A. (2000). Developmental trajectories of aggressive and non-aggressive conduct problems. *Journal of Quantitative Criminology*, *16*(2), 199–221.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, 110–133.

McGue, M., & Bouchard, T. J. (1984). Adjustment of twin data for the effects of age and sex. *Behavior genetics*, *14*(4), 325-343.

Miller, S., Malone, P. S., Dodge, K. A., & Conduct Problems Prevention Research Group. (2010). Developmental trajectories of boys' and girls' delinquency: Sex differences and links to later adolescent outcomes. *Journal of abnormal child psychology*, *38*, 1021-1032.

Moffitt, T. E. (1993). *Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy*. *100*(4), 674–701.

Moffitt, T. E. (2003). Life-course-persistent and adolescence-limited antisocial behavior: A 10-year research review and a research agenda.

Moffitt, T. E., Caspi, A., Harrington, H., & Milne, B. J. (2002). Males on the life-course-persistent and adolescence-limited antisocial pathways: Follow-up at age 26 years. *Development and psychopathology*, *14*(1), 179-207.

Monahan, K. C., Steinberg, L., & Cauffman, E. (2009). Affiliation with antisocial peers, susceptibility to peer influence, and antisocial behavior during the transition to adulthood. *Developmental psychology*, *45*(6), 1520.

Mrug, S., Elliott, M., Gilliland, M. J., Grunbaum, J. A., Tortolero, S. R., Cuccaro, P., & Schuster, M. (2008). Positive parenting and early puberty in girls: protective effects against aggressive behavior. *Archives of pediatrics & adolescent medicine*, *162*(8), 781-786.

Muris, P., Meesters, C., & van den Berg, F. (2003). The strengths and difficulties questionnaire (SDQ). *European child & adolescent psychiatry*, *12*(1), 1-8.

Murray, A. L., Eisner, M., & Ribeaud, D. (2020). Within-person analysis of developmental cascades between externalising and internalising problems. *Journal of Child Psychology and Psychiatry*, *61*(6), 681–688.

Muthén, B. (2004). Latent variable analysis. *The Sage handbook of quantitative methodology for the social sciences*, *345*(368), 106-109.

Muthén, B., & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. *Mplus web notes*, *4*(5), 1-22.

Muthén, B.O., du Toit, S.H.C., & Spisic, D. (1997). Robust inference using weighted-least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. *Psychometrika.*

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research*, *24*(6), 882-891.

Muthén, L. K., and B. O. Muthén. 1998-2019. Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Nagin, D. S. (1999). Analyzing Developmental Trajectories: A Semiparametric, Group-Based Approach. *Psychological Methods*, *4*(2), 139–157.

Nagin, D. (2005). *Group-based modeling of development*. Harvard University Press.

Nagin, D. S., Barker, T., Lacourse, E., & Tremblay, R. E. (2009). The inter-relationship of temporally distinct risk markers and the transition from childhood physical aggression to adolescent violent delinquency. *Monatsschrift für Kriminologie und Strafrechtsreform*, *92*(2-3), 102-116.

Nagin, D. S., & Tremblay, R. E. (2005). Developmental trajectory groups: Fact or a useful statistical fiction?. *Criminology*, *43*(4), 873-904.

Narusyte, J., Neiderhiser, J. M., Andershed, A. K., D'Onofrio, B. M., Reiss, D., Spotts, E., ... & Lichtenstein, P. (2011). Parental criticism and externalizing behavior problems in adolescents: The role of environment and genotype–environment correlation. *Journal of abnormal psychology*, *120*(2), 365.

Neale, M. C., & Cardon, L. R. (1992). NATO ASI series D: Behavioural and social sciences, Vol. 67. Methodology for genetic studies of twins and families.

NICHD Early Child Care Research Network, & Arsenio, W. F. (2004). Trajectories of physical aggression from toddlerhood to middle childhood: Predictors, correlates, and outcomes. *Monographs of the Society for Research in Child Development*, i–143.

Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., ... & Caspi, A. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and psychopathology*, *20*(2), 673.

Olson, S. L., Sameroff, A. J., Lansford, J. E., Sexton, H., Davis-Kean, P., Bates, J. E., ... & Dodge, K. A. (2013). Deconstructing the externalizing spectrum: Growth patterns of overt aggression, covert aggression, oppositional behavior, impulsivity/inattention, and emotion dysregulation between school entry and early adolescence. *Development and psychopathology*, *25*(3), 817-842.

Padilla-Walker, L. M., Memmott-Elison, M. K., & Coyne, S. M. (2018). Associations between Prosocial and Problem Behavior from Early to Late Adolescence. *Journal of Youth and Adolescence*, *47*(5), 961–975.

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

Peterson, R. D., & Krivo, L. J. (2009). Segregated spatial locations, race-ethnic composition, and neighborhood violent crime. *The Annals of the American Academy of Political and Social Science*, *623*(1), 93-107.

Pingault, J. B., Rijsdijk, F., Zheng, Y., Plomin, R., & Viding, E. (2015). Developmentally dynamic genome: Evidence of genetic influences on increases and decreases in conduct problems from early childhood to adolescence. *Scientific reports*, *5*(1), 1-9.

Piquero, N. L., Gover, A. R., MacDonald, J. M., & Piquero, A. R. (2005). The influence of delinquent peers on delinquency: Does gender matter?. *Youth & Society*, *36*(3), 251-275.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another?. *Behavioral and brain Sciences*, *10*(1), 1-16.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological bulletin*, *84*(2), 309.

Porsch, R. M., Middeldorp, C. M., Cherny, S. S., Krapohl, E., Van Beijsterveldt, C. E., Loukola, A., ... & Bartels, M. (2016). Longitudinal heritability of childhood aggression. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *171*(5), 697-707.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage Publications.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental review*, *41*, 71-90.

Raftery, A. E. (1995). Bayesian model selection in social research. Sociological methodology, 111-163.

Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., ... & Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and behavioral Disorders*, *15*(3), 130-142.

Rhee, S. H., & Waldman, I. D. (2002). Genetic and environmental influences on antisocial behavior: a meta-analysis of twin and adoption studies. *Psychological bulletin*, *128*(3), 490.

Rhemtulla, M., & Hancock, G. R. (2016). Planned missing data designs in educational psychology research. *Educational Psychologist*, *51*(3-4), 305-316.

Sacco, P., Bright, C. L., Jun, H.-J., & Stapleton, L. M. (2015). Developmental relations between alcohol and aggressive behavior among adolescents: Neighborhood and sociodemographic correlates. *International Journal of Mental Health and Addiction*, *13*(5), 603–617.

Samek, D. R., Goodman, R. J., Erath, S. A., McGue, M., & Iacono, W. G. (2016). Antisocial peer affiliation and externalizing disorders in the transition from adolescence to young adulthood: Selection versus socialization effects. *Developmental psychology*, *52*(5), 813.

Schaeffer, C. M., Petras, H., Ialongo, N., Masyn, K. E., Hubbard, S., Poduska, J., & Kellam, S. (2006). A comparison of girls' and boys' aggressive-disruptive behavior trajectories across elementary school: prediction to young adult antisocial outcomes. *Journal of consulting and clinical psychology*, *74*(3), 500.

Schaeffer, C. M., Petras, H., Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behavior across elementary school: links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental psychology*, *39*(6), 1020.

Schwartz, J. A., Solomon, S. J., & Valgardson, B. A. (2019). Socialization, selection, or both? The role of gene–environment interplay in the association between exposure to antisocial peers and delinquency. *Journal of quantitative criminology*, *35*, 1-26.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, *52*(3), 333-343.

Singh, G. K. (2003). Area deprivation and widening inequalities in US mortality, 1969–1998. *American journal of public health*, *93*(7), 1137-1143

Snyder, J. J., Schrepferman, L. P., Bullard, L., McEachern, A. D., & Patterson, G. R. (2012). Covert antisocial behavior, peer deviancy training, parenting processes, and sex differences in the development of antisocial behavior during childhood. *Development and psychopathology*, *24*(3), 1117-1138.

Snyder, H. R., Young, J. F., & Hankin, B. L. (2019). Chronic stress exposure and generation are related to the p-factor and externalizing specific psychopathology in youth. *Journal of Clinical Child & Adolescent Psychology*, *48*(2), 306-315.

Spano, R., Rivera, C., & Bolland, J. M. (2010). Are chronic exposure to violence and chronic violent behavior closely related developmental processes during adolescence? *Criminal Justice and Behavior*, *37*(10), 1160–1179.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual differences*, *42*(5), 893-898.

Sterba, S. K. (2014). Fitting nonlinear latent growth curve models with individually varying time points. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 630-647.

Thaweekoon, T. (2006). Effects of exposure to community violence on adolescent adjustment problems [PhD Thesis]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering*.

Tung, I., & Lee, S. S. (2017). Latent trajectories of adolescent antisocial behavior: Serotonin transporter linked polymorphic region (5-HTTLPR) genotype influences sensitivity to perceived parental support. *Development and Psychopathology*, *29*(1), 185-201.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current directions in psychological science*, *9*(5), 160-164.

Turkheimer, E., & Gottesman, I. I. (1991). Individual differences and the canalization of human behavior. Developmental Psychology, 27, 18-22.

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socioeconomic status: gene–environment interaction. *Journal of Child Psychology and Psychiatry*, *47*(7), 734-743.

Tuvblad, C., Narusyte, J., Grann, M., Sarnecki, J., & Lichtenstein, P. (2011). The genetic and environmental etiology of antisocial behavior from childhood to emerging adulthood. *Behavior genetics*, *41*(5), 629-640.

van Beijsterveldt, C. E. M., Bartels, M., Hudziak, J. J., & Boomsma, D. I. (2003). Causes of stability of aggression from early childhood to adolescence: A longitudinal genetic analysis in Dutch twins. *Behavior Genetics*, *33*(5), 591-605.

van Lier, P. A. C., Vitaro, F., Barker, E. D., Koot, H. M., & Tremblay, R. E. (2009). Developmental links between trajectories of physical violence, vandalism, theft, and alcohol-drug use from childhood to adolescence. *Journal of Abnormal Child Psychology*, *37*(4), 481–492.

Walden, B., McGue, M., Burt, S. A., & Elkins, I. (2004). Identifying shared environmental contributions to early substance use: the respective roles of peers and parents. *Journal of abnormal psychology*, *113*(3), 440.

Walters, G. D. (2011). The latent structure of life-course-persistent antisocial behavior: Is Moffitt's developmental taxonomy a true taxonomy?. *Journal of Consulting and Clinical Psychology*, *79*(1), 96.

Walters, G. D., & Ruscio, J. (2013). Trajectories of youthful antisocial behavior: Categories or continua? *Journal of Abnormal Child Psychology*, *41*(4), 653–666.

Wang, M. T., & Dishion, T. J. (2012). The trajectories of adolescents' perceptions of school climate, deviant peer affiliation, and behavioral problems during the middle school years. *Journal of research on adolescence*, *22*(1), 40-53.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child development perspectives*, *4*(1), 10-18.

Windle, M. (2000). A latent growth curve model of delinquent activity among adolescents. *Applied Developmental Science*, *4*(4), 193-207.

Xu, Y., Zhou, Y., Zhao, J., Xuan, Z., Li, W., Han, L., & Liu, H. (2021). The relationship between shyness and aggression in late childhood: The multiple mediation effects of parent-child conflict and self-control. *Personality and individual differences*, *182*, 111058.

Yan, J., Schoppe-Sullivan, S. J., & Beauchaine, T. P. (2021). Paternal antisociality and growth in child delinquent behaviors: Moderating effects of child sex and respiratory sinus arrhythmia. *Developmental psychobiology*, *63*(5), 1466-1481.

Yoon, D., Yoon, S., Yoon, M., & Snyder, S. M. (2019). Developmental trajectories of deviant peer affiliation in adolescence: Associations with types of child maltreatment and substance use. *Children and Youth Services Review*, *105*, 104446.

Yuan, K. H., & Chan, W. (2016). Measurement invariance via multigroup SEM: Issues and solutions with chi-square-difference tests. *Psychological methods*, *21*(3), 405.

APPENDIX A: SUPPLEMENTAL MATERIAL FOR CHAPTER 1

Tables

Table S1.1. *Standardized variance estimates from univariate ACE model for CBCL Conduct Problems scale*

	Variance estimates		
	a	c	e
Univariate			
Intercept	**.50***	**.20***	**.31***
Slope	**.54***	.11	**.35***

Note: Bold font and asterisk indicate that the estimate was significantly different than zero at p<.05.

Figures

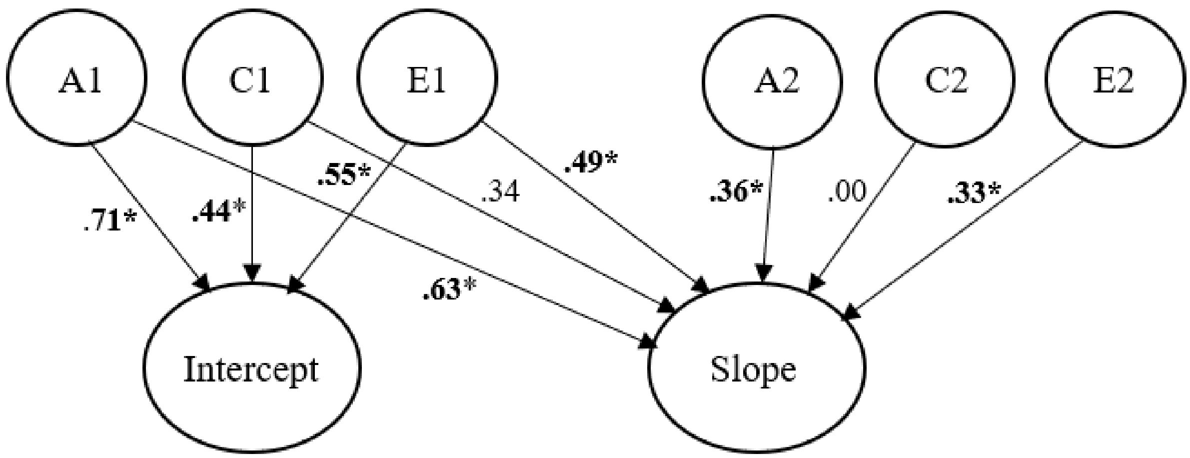


Figure S1.1. *Path diagram of a bivariate twin model for CBCL Conduct Problems scale.* The variance in the intercept and the slope is partitioned into additive genetic effects (A1 and A2), shared environmental effects (C1 and C2), and nonshared environmental effects (E1 and E2). For ease of presentation, this path diagram represents one twin in a pair. Standardized path estimates are squared to represent the proportion of variance accounted for.

APPENDIX B: SUPPLEMENTAL MATERIAL FOR CHAPTER 2

Supplementary Methods

Measures

AGG

Teacher reports of AGG (20 items) were obtained at the first two assessment waves via the Achenbach Teacher Report Form (TRF; Achenbach & Rescorla, 2001). Teachers rated the twins' behaviors during the preceding six months using a three-point scale (0=never to 2=often/mostly true). The teachers of 115 participants were not available for assessment (because the twins were home-schooled or because parental consents to contact the teachers were completed incorrectly, etc.). Our teacher participation rate across both subsamples was 86% at Wave 1 and 60% at Wave 2, with teacher reports available for 1,551 and 453 participants, respectively. In addition, the twins completed the Youth Self-Report (YSR) at Waves 2 and 3, reporting on their own AGG (17 items) during the preceding six months using the three-point scale described above (Achenbach & Rescorla, 2001). Reports were available from 99% and 83% of the twins at Waves 2 and 3, respectively.

Data Analytic Strategy

Item Response Theory (IRT) Analyses

Analyses were conducted in M*plus* 8.4 (Muthén & Muthén, 1998-2019) using the weighted least squares mean- and variance-adjusted (WLSMV) estimator. As the twins were nested within families, analyses accounted for the nonindependence of observations using the CLUSTER command. We made use of an IRT model, which relates an individual's trait level (θ) to their performance on a series of items while accounting for item characteristics. Given the relatively low frequency of AGG in the sample, and consistent with prior IRT analyses of antisocial behavior in the TBED-C (Burt, Donnellan, Slawinski, & Klump, 2016), items were coded dichotomously, with responses of 1 or 2 on the ASEBA instruments collapsed into a single "present" category (i.e., 0=behavior absent, 1=behavior present). We then applied the two-parameter logistic (2PL) model, which estimates two item characteristics: discrimination (α) and difficulty (β) (Embretson & Reise, 2013). Item discrimination indicates the extent to which endorsement of a given item relates to one's trait level; the higher the discrimination, the more accurately the item assesses standing on the trait of interest. Item difficulty indicates the trait level needed to have a 50% chance of endorsing the item. A higher difficulty thus indicates that one would need a relatively high standing on the latent trait continuum to receive a 1. Such an item would be diagnostic of high levels of AGG.

Initial models included all AGG items from each available informant (i.e., 18 items from the CBCL and 20 items from the TRF at Wave 1; 18 items from the CBCL, 20 from the TRF, and 17 from the YSR at Wave 2; 18 from the CBCL and 17 from the YSR at Wave 3). Each wave comprised one latent factor. Items were omitted from subsequent models based on their characteristics (i.e., low discrimination and/or low endorsement rate) as well as theoretical considerations (i.e., whether they assessed emotional dysregulation or verbal AGG rather than physical AGG). Model fit was evaluated with three indices: root mean square error of approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI). RMSEA values below .06 and CFI and TLI values above .95 are considered to indicate good fit (Hu & Bentler, 1999). For each model, we examined modification indices to identify sources of misfit and added the suggested residual covariances between items to the model. AGG mean scores were computed based on participants' responses to the items retained in the final IRT model.

Supplementary Results

IRT Modeling

The initial IRT model at each wave included all ASEBA AGG items from all available informants. The second model omitted items assessing emotion dysregulation (e.g., sudden changes in mood) and/or verbal AGG (e.g., argues). The final model excluded these items as well as those with standardized discrimination values < .3 (Hair, Anderson, Tatham, & Black, 1998; i.e., suspicious) or endorsement rates near zero (i.e., threatens others). Fit indices are reported in Table S2.2, and parameter estimates from the first two models are reported in Tables S2.3 and S2.4, respectively. The final model at each wave retained eight items from each available informant. Fit was good at Waves 1 and 3, according to all fit indices. At Wave 2, the CFI and TLI indicated some degree of misfit, even after residual covariances were added, but the RMSEA was well below recommended cutoffs (Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999; Steiger, 2007), indicating good fit. Discrimination and difficulty parameter estimates from the final model are reported in Table S2.5. As expected, items assessing behaviors that could be considered illegal (e.g., destroys things belonging to others, physically attacks people) were more difficult to endorse than were those assessing less extreme behaviors (e.g., disobedient at home, hot temper). Teacher-reported items generally discriminated better than did other informants' and were more difficult to endorse at Wave 1 than parent-reported items, consistent with prior work identifying somewhat higher rates of psychopathology in parent reports relative to teacher reports (Ferdinand et al., 2003). Mean scores were computed based on the items retained in the final model. Scores were multiplied by 24 to correspond to the highest number of items included at any wave (i.e., Wave 2) and then log-transformed to account for positive skew.

Tables

Table S2.1. *Youth AGG by age*

Age	Total *N*	Mean AGG (SD)	Range AGG
6	598	1.84 (2.06)	0-8
7	410	1.53 (1.71)	0-7
8	346	1.74 (1.98)	0-8
9	342	1.22 (1.61)	0-8
10	352	1.34 (1.78)	0-8
11	119	1.01 (1.38)	0-7
12	82	.73 (1.22)	0-7
13	92	.73 (1.29)	0-5
14	185	.77 (1.42)	0-8
15	234	.59 (1.12)	0-6
16	164	.44 (.99)	0-6
17	136	.69 (1.13)	0-6
18	68	.53 (1.09)	0-5
19	36	.28 (.70)	0-3
20-21	7	.00 (.00)	0-0

Note: AGG sum scores were computed based on maternal reports on eight CBCL items.

Table S2.2. *IRT model fit*

		RMSEA (90% CI)	CFI	TLI
Wave 1				
	Model 1	.010 (.007-.013)	.996	.994
	Model 2	.032 (.028-.036)	.977	.961
	Model 3	.013 (.004-.019)	.997	.996
Wave 2				
	Model 1	.029 (.027-.031)	.841	.831
	Model 2	.029 (.025-.033)	.901	.890
	Model 3	.032 (.027-.036)	.905	.890
Wave 3				
	Model 1	.034 (.029-.040)	.912	.906
	Model 2	.036 (.026-.045)	.947	.940
	Model 3	.036 (.024-.048)	.960	.952

Note: Model 1 contained all ASEBA AGG items from each available informant at a given wave. Model 2 contained all items assessing physical AGG. Model 3 was the final model, containing items that assessed physical AGG and had good psychometric properties.

Table S2.3. *Standardized IRT parameter estimates from models containing all ASEBA AGG items*

	Wave 1					
	CBCL		TRF		YSR	
	α	*β*	*A*	*β*	*α*	*β*
Argues	.22	-.28	.91	.61	-	-
Defiant, talks back	-	-	.84	1.11	-	-
Cruelty, bullying, or meanness	.34	1.06	.82	1.14	-	-
Demands attention	.16	.41	.71	.70	-	-
Destroys own things	.47	.96	.82	1.52	-	-
Destroys things belonging to others	.45	.90	.84	1.72	-	-
Disobedient at home*	.36	.00	-	-	-	-
Disobedient at school	.69	.96	.91	1.00	-	-
Many fights	.28	1.41	.81	1.60	-	-
Physically attacks people	.31	1.04	.86	1.61	-	-
Screams	.29	.82	.80	1.86	-	-
Explosive and unpredictable	-	-	.87	1.51	-	-
Easily frustrated	-	-	.76	1.09	-	-
Stubborn, sullen, or irritable	.24	.20	.87	1.02	-	-
Sudden changes in mood	.30	.61	.74	1.20	-	-
Sulks	.18	1.03	.72	1.10	-	-
Suspicious	.39	1.62	.81	1.86	-	-
Teases	.19	.91	.72	1.42	-	-
Hot temper	.25	.27	.87	1.41	-	-
Threatens people	.32	1.49	.88	1.87	-	-
Unusually loud	.22	.72	.70	1.29	-	-

Table S2.3 (cont'd)

	Wave 2					
	CBCL		**TRF**		**YSR**	
	α	*β*	*A*	*β*	*α*	*β*
Argues	.43	-.32	.75	-.07	.25	-.42
Defiant, talks back	-	-	.93	.02	-	-
Cruelty, bullying, or meanness	.61	.92	.89	.18	.44	.35
Demands attention	.50	.38	.75	.21	.21	.60
Destroys own things	.59	.91	.71	.94	.53	.83
Destroys things belonging to others	.69	.67	.78	.95	.62	1.18
Disobedient at home*	.55	.00	-	-	.35	.19
Disobedient at school	.71	.61	.89	.19	.57	.70
Many fights	.66	1.12	.87	.67	.52	1.08
Physically attacks people	.64	1.03	.85	.87	.54	1.18
Screams	.54	.70	.84	1.07	.52	.62
Explosive and unpredictable	-	-	.96	.46	-	-
Easily frustrated	-	-	.83	.33	-	-
Stubborn, sullen, or irritable	.39	.01	.72	.16	.001	-.11
Sudden changes in mood	.42	.22	.81	.26	.28	-.20
Sulks	.44	.60	.72	.46	-	-
Suspicious	.44	1.06	.86	.72	.30	.36
Teases	.45	.50	.76	.39	.20	.59
Hot temper	.47	.27	.92	.53	.44	-.01
Threatens people	.74	1.07	1.00	.69	.52	1.08
Unusually loud	.53	.42	.83	.49	.36	-.04

Table S2.3 (cont'd)

	Wave 3					
	CBCL		TRF		YSR	
	α	*β*	*A*	*β*	*α*	*β*
Argues	.79	-.42	-	-	.41	-.41
Defiant, talks back	-	-	-	-	-	-
Cruelty, bullying, or meanness	.86	.87	-	-	.55	.49
Demands attention	.64	.52	-	-	.32	.71
Destroys own things	.67	1.20	-	-	.40	1.39
Destroys things belonging to others	.67	1.05	-	-	.63	1.58
Disobedient at home*	.85	.00	-	-	.57	.23
Disobedient at school	.64	.94	-	-	.44	1.10
Many fights	.70	1.74	-	-	.46	1.55
Physically attacks people	.86	1.15	-	-	.70	1.71
Screams	.71	.95	-	-	.28	1.13
Explosive and unpredictable	-	-	-	-	-	-
Easily frustrated	-	-	-	-	-	-
Stubborn, sullen, or irritable	.71	-.10	-	-	.34	-.41
Sudden changes in mood	.77	.04	-	-	.53	-.13
Sulks	.65	.72	-	-	-	-
Suspicious	.45	1.13	-	-	.32	.53
Teases	.46	.66	-	-	.31	.54
Hot temper	.85	.21	-	-	.59	.07
Threatens people	.86	1.26	-	-	.35	1.72
Unusually loud	.47	.97	-	-	.38	.24

Note: *The CBCL item "disobedient at home" was used as the anchor item at each wave for model identification purposes.

Table S2.4. *Standardized IRT parameter estimates from models containing all ASEBA items assessing physical AGG*

	Wave 1					
	CBCL		TRF		YSR	
	α	*β*	*A*	*β*	*α*	*β*
Cruelty, bullying, or meanness	.59	1.08	.88	1.24	-	-
Destroys own things	.78	.99	.70	1.64	-	-
Destroys things belonging to others	.88	.91	.75	1.84	-	-
Disobedient at home*	.75	.00	-	-	-	-
Disobedient at school	.69	1.05	.96	1.12	-	-
Many fights	.65	1.40	.82	1.71	-	-
Physically attacks people	.64	1.05	.87	1.72	-	-
Explosive and unpredictable	-	-	.83	1.63	-	-
Suspicious	.61	1.65	.73	1.98	-	-
Hot temper	.58	.26	.76	1.54	-	-
Threatens people	.64	1.49	.85	1.99	-	-
	Wave 2					
	CBCL		**TRF**		**YSR**	
	α	*B*	*A*	*β*	*α*	*β*
Cruelty, bullying, or meanness	.58	.91	.83	.18	.50	.23
Destroys own things	.62	.81	.66	.94	.53	.77
Destroys things belonging to others	.67	.63	.76	.91	.61	1.14
Disobedient at home*	.52	.00	-	-	.30	.23
Disobedient at school	.74	.50	.87	.14	.64	.55
Many fights	.71	1.00	.90	.55	.58	.95

Table S2.4 (cont'd)

Physically attacks people	.67	.94	.88	.76	.60	1.05
Explosive and unpredictable	-	-	.96	.38	-	-
Suspicious	.37	1.12	.87	.63	.30	.33
Hot temper	.51	.18	.94	.43	.46	-.08
Threatens people	.75	1.00	1.01	.58	.53	1.02
			Wave 3			
	CBCL		**TRF**		**YSR**	
	α	*β*	*A*	*β*	*α*	*B*
Cruelty, bullying, or meanness	.78	.92	-	-	.59	.44
Destroys own things	.71	1.14	-	-	.40	1.38
Destroys things belonging to others	.70	.99	-	-	.64	1.56
Disobedient at home*	.83	.00	-	-	.62	.16
Disobedient at school	.66	.91	-	-	.44	1.09
Many fights	.80	1.62	-	-	.50	1.49
Physically attacks people	.91	1.08	-	-	.71	1.67
Explosive and unpredictable	-	-	-	-	-	-
Suspicious	.20	1.36	-	-	.28	.56
Hot temper	.78	.25	-	-	.65	-.01
Threatens people	.94	1.15	-	-	.40	1.66

Note: *The CBCL item "disobedient at home" was used as the anchor item at each wave for model identification purposes.

Table S2.5. *Standardized IRT parameter estimates from final model*

	Wave 1					
	CBCL		TRF		YSR	
	α	*β*	*A*	*β*	*α*	*β*
Cruelty, bullying, or meanness	.36	1.06	.83	1.14	-	-
Destroys own things	.55	.94	.80	1.53	-	-
Destroys things belonging to others	.46	.91	.84	1.72	-	-
Disobedient at home*	.37	.00	-	-	-	-
Disobedient at school	.72	.96	.93	1.00	-	-
Many fights	.35	1.39	.86	1.60	-	-
Physically attacks people	.33	1.04	.91	1.60	-	-
Explosive and unpredictable	-	-	.80	1.54	-	-
Hot temper	.25	.27	.79	1.44	-	-
	Wave 2					
	CBCL		**TRF**		**YSR**	
	α	*β*	*A*	*β*	*α*	*B*
Cruelty, bullying, or meanness	.52	1.00	.80	.23	.44	.32
Destroys own things	.63	.80	.64	.98	.55	.75
Destroys things belonging to others	.66	.65	.74	.94	.65	1.09
Disobedient at home*	.52	.00	-	-	.33	.20
Disobedient at school	.77	.47	.86	.17	.59	.62
Many fights	.74	.97	.92	.54	.60	.94
Physically attacks people	.68	.92	.88	.76	.57	1.10
Explosive and unpredictable	-	-	.96	.39	-	-
Hot temper	.45	.27	.93	.45	.44	-.04

Table S2.5 (cont'd)

	Wave 3					
	CBCL		**TRF**		**YSR**	
	α	*β*	*A*	*B*	*α*	*B*
Cruelty, bullying, or meanness	.78	.91	-	-	.59	.43
Destroys own things	.75	1.10	-	-	.42	1.36
Destroys things belonging to others	.73	.96	-	-	.72	1.46
Disobedient at home*	.82	.00	-	-	.50	.28
Disobedient at school	.70	.86	-	-	.57	.95
Many fights	.82	1.59	-	-	.44	1.56
Physically attacks people	.92	1.06	-	-	.64	1.75
Explosive and unpredictable	-	-	-	-	-	-
Hot temper	.71	.32	-	-	.67	-.03

Note: *The CBCL item "disobedient at home" was used as the anchor item at each wave for model identification purposes.

Table S2.6. *LGM fit statistics and parameter estimates for multi-informant mean AGG scores*

Unconditional LGM model fit statistics					
Model	-2lnL	χ^2 (df)	AIC	BIC	SABIC
Linear growth	**8049.46**	**-**	**8061.46**	**8095.25**	**8076.18**
Means model	8189.34	139.88† (3)	8195.34	8212.23	8202.70

Conditional LGM parameter estimates			
Parameter	Estimate	S.E.	p-value
Intercept			
Mean	**.938**	**.076**	**<.001**
Variance	**.510**	**.046**	**<.001**
ADI	**.084**	**.031**	**.007**
Income	**-.094**	**.034**	**.006**
Conflict	**.335**	**.026**	**<.001**
Nurturance	.001	.024	.977
Sex	**.317**	**.053**	**<.001**
Race	.032	.079	.684
Slope			
Mean	-.021	.011	.060
Variance	**.005**	**.001**	**<.001**
ADI	-.003	.005	.567
Income	-.002	.005	.698
Conflict	**-.018**	**.004**	**<.001**
Nurturance	.003	.004	.445
Sex	**-.018**	**.008**	**.031**
Race	-.007	.011	.570
Covariances			
Intercept with slope	**-.036**	**.006**	**<.001**
Intercept 1 with Intercept 2			
MZ	**.481**	**.048**	**<.001**
DZ	**.341**	**.046**	**<.001**
Slope 1 with Slope 2			
MZ	**.006**	**.001**	**<.001**
DZ	**.003**	**.001**	**.001**
Intercept 1 with Slope 2			
MZ	**-.040**	**.006**	**<.001**
DZ	**-.030**	**.006**	**<.001**
Residual variance	**.319**	**.024**	**<.001**

Note: Bold font indicates $p<.05$. †Significant change in chi-square at $p<.05$.

Table S2.7. *LCGA fit statistics and predictors of group membership for multi-informant mean AGG scores*

Unconditional LCGA model fit statistics			
	AIC	BIC	Smallest group (% of sample)
# of groups			
1	8292.06	8308.95	100
2	7893.68	7927.46	45.6
3	7797.62	7848.29	12.6
4	7713.73	7781.29	3.3
5	7676.40	7760.85	2.7
6	7639.27	7740.61	2.9
Conditional LCGA parameter estimates from 3-group model			
Persistent v. low/stable			
	Estimate	S.E.	*p*-value
ADI	**.403**	**.114**	**<.001**
Income	**-.393**	**.090**	**<.001**
Conflict	**1.233**	**.117**	**<.001**
Nurturance	.117	.095	.216
Sex	**.808**	**.201**	**<.001**
Race	.007	.207	.971
Persistent v. desisting			
	Estimate	S.E.	*p*-value
ADI	**.350**	**.162**	**.031**
Income	**-.449**	**.134**	**.001**
Conflict	**.352**	**.139**	**.011**
Nurturance	.151	.134	.258
Sex	.142	.308	.645
Race	-.283	.317	.372
Desisting v. low/stable			
	Estimate	S.E.	*p*-value
ADI	.054	.104	.607
Income	.056	.123	.646
Conflict	**.881**	**.106**	**<.001**
Nurturance	-.034	.098	.728
Sex	**.666**	**.195**	**.001**
Race	.290	.257	.258

Note: Bold font indicates *p*<.05.

Table S2.8. *Parameter estimates from conditional LGM with data in "long" format*

Parameter	Estimate	S.E.	p-value
		Intercept	
Mean	**.708**	**.047**	**<.001**
Variance	**.286**	**.025**	**<.001**
ADI	**.054**	**.019**	**.005**
Income	**-.071**	**.021**	**.001**
Conflict	**.291**	**.017**	**<.001**
Nurturance	.028	.017	.110
Sex	**.180**	**.033**	**<.001**
Race	.008	.048	.870
		Slope	
Mean	**-.047**	**.007**	**<.001**
Variance	**.002**	**.001**	**<.001**
ADI	-.003	.003	.394
Income	.003	.003	.380
Conflict	**-.019**	**.003**	**<.001**
Nurturance	.000	.003	.950
Sex	**-.017**	**.005**	**.001**
Race	.006	.007	.418
Intercept with slope	**-.020**	**.003**	**<.001**
Residual variance	**.149**	**.014**	**<.001**

Note: Parent-reported AGG is the outcome measure. Bold font indicates $p<.05$.

Table S2.9. *LCGA fit statistics and predictors of AGG group membership with 1 twin randomly selected per family (N = 1,030)*

Unconditional LCGA model fit statistics			
	AIC	BIC	Smallest group (% of sample)
# of groups			
1	2976.49	2991.28	100
2	2708.76	2738.35	38.7
3	2589.30	2633.69	15.0
4	2533.19	2592.38	9.4
5	2474.60	2548.59	2.0
6	2454.29	2543.07	2.7
Conditional LCGA parameter estimates from 3-group model			
Persistent v. low/stable			
	Estimate	S.E.	*p*-value
ADI	.281	.146	.055
Income	**-.394**	**.120**	**.001**
Conflict	**1.221**	**.161**	**<.001**
Nurturance	.158	.156	.313
Sex	**.684**	**.265**	**.010**
Race	.320	.340	.347
Persistent v. desisting			
	Estimate	S.E.	*p*-value
ADI	.180	.189	.340
Income	**-.358**	**.135**	**.008**
Conflict	**.467**	**.166**	**.005**
Nurturance	.169	.201	.401
Sex	.169	.335	.613
Race	.167	.418	.689
Desisting v. low/stable			
	Estimate	S.E.	*p*-value
ADI	.101	.120	.400
Income	-.036	.114	.753
Conflict	**.755**	**.115**	**<.001**
Nurturance	-.011	.117	.924
Sex	**.514**	**.204**	**.012**
Race	.153	.271	.573

Note: Parent-reported AGG is the outcome measure. Bold font indicates *p*<.05.

Figures

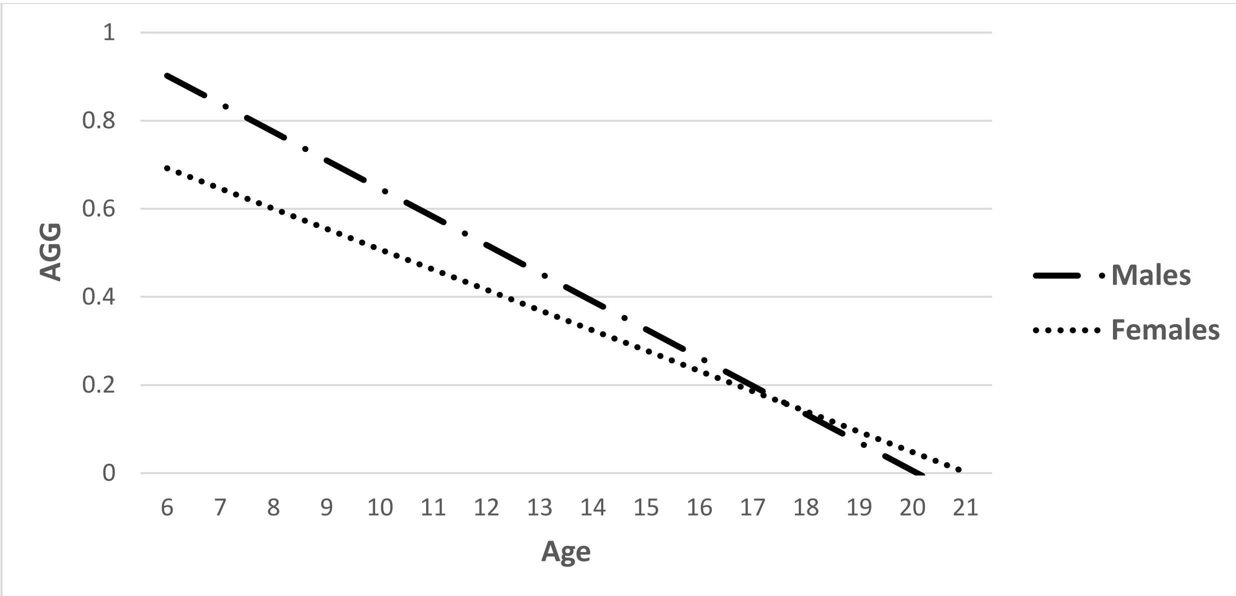


Figure S2.1. *Mean patterns of AGG for males and females from the latent growth curve model (variable-centered).*

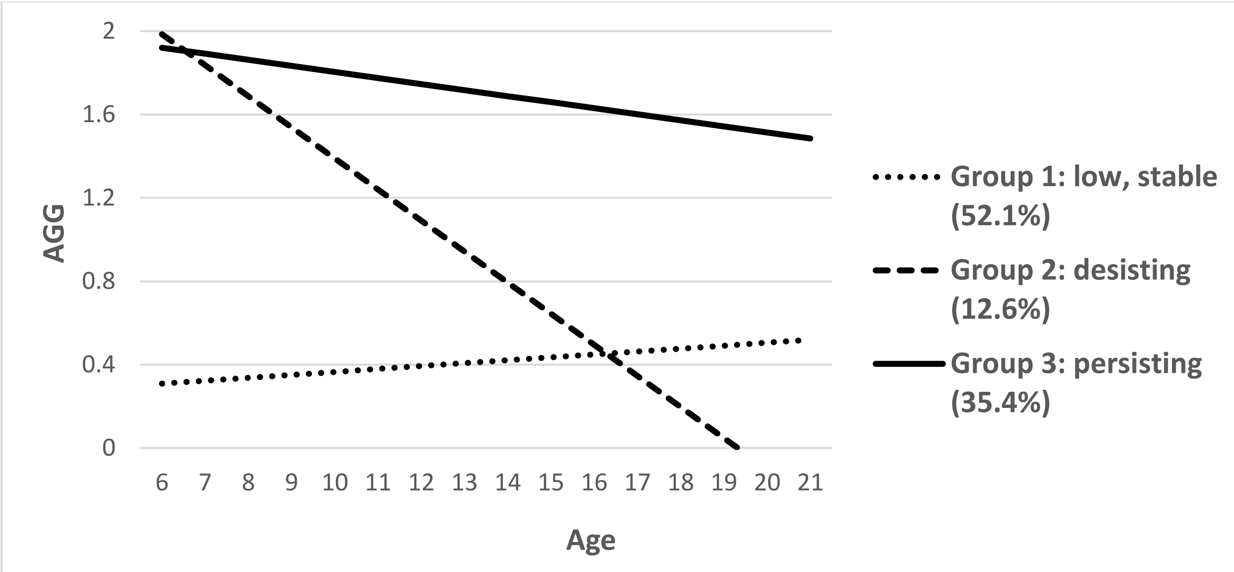


Figure S2.2. *Unconditional three-group model of AGG trajectories based on multi-informant mean scores.*

APPENDIX C: SUPPLEMENTAL MATERIAL FOR CHAPTER 3

Supplementary Methods

Data Analytic Strategy

Item Response Theory (IRT) Analyses

Analyses were conducted in M*plus* 8.4 (Muthén & Muthén, 1998-2019) using the weighted least squares mean- and variance-adjusted (WLSMV) estimator. As the twins were nested within families, analyses accounted for the nonindependence of observations using the CLUSTER command. We made use of an IRT model, which relates an individual's trait level (θ) to their performance on a series of items while accounting for item characteristics. Items were drawn from the Rule-Breaking (RB) scale on the ASEBA instruments, including the Child Behavior Checklist (CBCL), Teacher Report Form (TRF), and Youth Self-Report (YSR) (Achenbach & Rescorla, 2001). Given the relatively low frequency of RB in the sample, and consistent with prior IRT analyses of antisocial behavior in the TBED-C (Burt, Donnellan, Slawinski, & Klump, 2016), items were coded dichotomously, with responses of 1 or 2 on the ASEBA instruments collapsed into a single "present" category (i.e., 0=behavior absent, 1=behavior present). We then applied the two-parameter logistic (2PL) model, which estimates two item characteristics: discrimination (α) and difficulty (β) (Embretson & Reise, 2013). Item discrimination indicates the extent to which endorsement of a given item relates to one's trait level; the higher the discrimination, the more accurately the item assesses standing on the trait of interest. Item difficulty indicates the trait level needed to have a 50% chance of endorsing the item. A higher difficulty thus indicates that one would need a relatively high standing on the latent trait continuum to receive a 1. Such an item would be diagnostic of high levels of RB.

We omitted items assessing substance use and sexual problems, given the very low base rates of these behaviors in childhood. At Wave 1, for example, no informant indicated that drug use was a concern, and only one teacher endorsed the item "seems preoccupied with sex". Other behaviors that were endorsed at low rates during Wave 1 were not endorsed at all during later assessments; for instance, no caregiver endorsed “sets fires” at Wave 3 or “vandalism” at Waves 2 or 3. In addition, the item “runs away from home” was endorsed by less than 2% of caregivers and youth at Waves 2 and 3 (this item is not included on the TRF). Thus, initial models included 18 items at Waves 1 and 3 and 27 items at Wave 2 (i.e., 9 items from each available informant) assessing behaviors that were present to some degree at all assessments. Each wave comprised one latent factor. Items were omitted from the final model based on their characteristics (i.e., low discrimination and/or high difficulty). Model fit was evaluated with three indices: root mean square error of approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis Index (TLI). RMSEA values below .06 and CFI and TLI values above .95 are considered to indicate good fit (Hu & Bentler, 1999). For each model, we examined modification indices to identify sources of misfit and added the suggested residual covariances between items to the model. RB mean scores were computed based on participants’ responses to the items retained in the final IRT model.

Measurement Invariance Analyses

After selecting the items that functioned well at each assessment wave, we conducted a series of analyses to test for measurement invariance (i.e., the extent to which items assessed the same construct over time, a prerequisite for examining developmental change; Widaman, Ferrer, & Conger, 2010). Analyses made use of the weighted least squares mean and variance adjusted (WLSMV) estimator and delta parameterization in Mplus 8.4, an appropriate method for dichotomous outcome data (Múthen & Asparouhov, 2002). The initial model examined configural invariance, with all items from a given wave loading onto a single factor and item loadings and thresholds free to vary across waves. Factor means and variances were fixed to 0 and 1, respectively, and scale factors were fixed to 1, for model identification purposes. Modification indices were examined to identify sources of misfit, and the suggested residual covariances between items were added to the model. The second model tested for scalar invariance, with item loadings and thresholds constrained to equality within each informant over time. To identify the model, the mean and variance of the latent factor at Wave 1 were constrained to 0 and 1, respectively, and the scale factors were fixed to 1 for each informant at one time point (i.e., scale factors for CBCL and TRF items were fixed to 1 at Wave 1, and scale factors for YSR items were fixed to 1 at Wave 2). The same residual covariances between items that were included in the configural invariance model were also included when assessing scalar invariance. Model fit was assessed using the RMSEA, CFI, TLI, and χ^2 difference test (Múthen, du Toit, & Spisic, 1997). Minimal changes in fit between the two models would indicate that the items measured RB consistently across assessment waves. (Other models to assess measurement invariance, such as a metric invariant model that allows item thresholds, but not loadings, to vary across waves, are not identified when examining binary items (Múthen & Asparouhov, 2002). As a result, we focused on the respective fits of the configural and scalar invariance models.)

Supplementary Results

IRT Modeling

The initial IRT model at each wave included 9 ASEBA RB items from each available informant. We subsequently omitted the CBCL/TRF item "prefers being with older kids" and corresponding YSR item "prefer being with younger kids" due to their low discrimination values. Fit indices are reported in Table S3.2, and parameter estimates are reported in Tables S3.3 and S3.4. Fit was good at Wave 1 according to all fit indices. At Waves 2 and 3, the CFI and TLI indicated some degree of misfit, even after residual covariances were added, but the RMSEA was well below recommended cutoffs (Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999; Steiger, 2007), indicating good fit. The final model at each wave retained eight items from each available informant. As expected, items assessing illegal behaviors (e.g., steals) were more difficult to endorse than were those assessing relatively minor rule violations (e.g., lying or cheating). Some items (e.g., truancy) were much easier to endorse at later assessment waves relative to Wave 1, reflecting the greater frequency of those behaviors during adolescence.

Teacher-reported items generally discriminated best, whereas self-reported items had the lowest discrimination values. We computed mean RB scores based on the items retained in the final model. Scores were multiplied by 24 to correspond to the highest number of items included at any wave (i.e., Wave 2) and then log-transformed to account for positive skew.

Measurement Invariance

Model fit estimates from the measurement invariance analyses are shown in Table S3.5. The χ^2 difference test was significant, suggesting decrement in model fit. However, the χ^2 fit statistic is heavily influenced by sample size (Bentler & Bonett, 1980) (see Yuan & Chan (2016) for a discussion of limitations specific to χ^2 difference testing). Moreover, the absolute fit indices were nearly identical for the two models, meaning that both the item loadings and thresholds could be constrained to equality across waves (Putnick & Bornstein, 2016). The items thus appeared to assess RB in a way that was consistent and developmentally appropriate across assessment waves.

Tables

Table S3.1. *Youth RB by age*

Age	Total *N*	Mean RB (SD)	Range RB
6	599	2.80 (3.45)	0-21
7	414	2.48 (3.27)	0-18
8	347	2.74 (3.39)	0-21
9	349	2.14 (3.48)	0-21
10	354	2.42 (3.44)	0-21
11	119	2.31 (2.95)	0-15
12	82	2.14 (2.19)	0-10
13	92	2.48 (2.89)	0-12
14	186	2.67 (3.43)	0-23
15	234	2.72 (3.21)	0-15
16	170	2.87 (3.17)	0-21
17	136	3.69 (3.47)	0-15
18	68	2.81 (3.13)	0-16.5
19	36	3.24 (4.01)	0-18
20-21	7	4.50 (3.35)	0-9

Note: Multi-informant composite mean RB scores are the outcome measure.

Table S3.2. *IRT model fit*

		RMSEA (90% CI)	CFI	TLI
Wave 1				
	Model 1	.019 (.015-.024)	.981	.973
	Model 2	.021 (.016-.026)	.984	.975
Wave 2				
	Model 1	.028 (.024-.033)	.935	.925
	Model 2	.032 (.028-.037)	.936	.924
Wave 3				
	Model 1	.040 (.029-.050)	.917	.900
	Model 2	.044 (.032-.055)	.923	.905

Note: Model 1 contained 9 ASEBA RB items from each available informant at a given wave. Model 2 was the final model, containing items that were endorsed at all waves and had good psychometric properties.

Table S3.3. *Standardized IRT parameter estimates from Model 1 at each wave*

	CBCL		TRF		YSR	
Wave 1						
	α	*β*	*α*	*B*	*α*	*β*
Breaks rules*	.59	.00	.89	.64	-	-
Doesn't seem to feel guilty after misbehaving	.51	.69	.75	.87	-	-
Hangs around with others who get in trouble	.50	.99	.72	.80	-	-
Lies or cheats	.60	.31	.82	.96		
Prefers being with older kids	.18	.35	.27	1.48	-	-
Prefers being with younger kids	-	-	-	-	-	-
Steals at home	.63	1.25	-	-	-	-
Steals outside the home	.65	1.53	-	-	-	-
Steals	-	-	.75	1.59	-	-
Swears	.61	1.06	.84	1.76	-	-
Truant	.54	2.24	.29	1.86	-	-
Tardy to school or class	-	-	.34	1.19	-	-
Wave 2						
	CBCL		**TRF**		**YSR**	
	α	*β*	*α*	*B*	*α*	*β*
Breaks rules*	.62	.00	.94	-.30	.35	.28
Doesn't seem to feel guilty after misbehaving	.54	.16	.93	-.35	.37	.27
Hangs around with others who get in trouble	.69	.30	.81	-.32	.39	.36
Lies or cheats	.51	.25	.91	-.02	.19	.81
Prefers being with older kids	.36	.30	.29	.85	-	-
Prefers being with younger kids	-	-	-	-	-.02	.81
Steals at home	.73	.62	-	-	.51	.97
Steals outside the home	.87	.70	-	-	.60	1.21
Steals	-	-	.78	.77	-	-
Swears	.47	.15	.79	.08	.23	-.30
Truant	.88	.68	.67	.21	.63	.72
Tardy to school or class	-	-	.66	-.07	-	-
Wave 3						
	CBCL		**TRF**		**YSR**	
	α	*β*	*α*	*B*	*α*	*β*
Breaks rules*	.76	.00	-	-	.63	.16
Doesn't seem to feel guilty after misbehaving	.66	.19	-	-	.34	.54

Table S3.3 (cont'd)

Hangs around with others who get in trouble	.83	.32	-	-	.49	.34
Lies or cheats	.69	-.06	-	-	.18	.90
Prefers being with older kids	.31	.53	-	-	-	-
Prefers being with younger kids	-	-	-	-	.37	-.47
Steals at home	.95	.91	-	-	.50	1.26
Steals outside the home	.91	.88	-	-	.76	.88
Steals	-	-	-	-	-	-
Swears	.57	.09	-	-	.58	-.89
Truant	.76	.63	-	-	.54	.62
Tardy to school or class	-	-	-	-	-	-

Note: *The CBCL item "breaks rules" was used as the anchor item at each wave for model identification purposes.

Table S3.4. *Standardized IRT parameter estimates from Model 2 at each wave*

	Wave 1					
	CBCL		**TRF**		**YSR**	
	α	*β*	*α*	*β*	*α*	*β*
Breaks rules*	.57	.00	.88	.63	-	-
Doesn't seem to feel guilty after misbehaving	.50	.68	.84	.80	-	-
Hangs around with others who get in trouble	.50	.98	.71	.80	-	-
Lies or cheats	.59	.31	.85	.92		
Steals at home	.64	1.24	-	-	-	-
Steals outside the home	.67	1.50	-	-	-	-
Steals	-	-	.75	1.58	-	-
Swears	.60	1.06	.83	1.76	-	-
Truant	.54	2.23	.29	1.85	-	-
Tardy to school or class	-	-	.34	1.19	-	-
	Wave 2					
	CBCL		**TRF**		**YSR**	
	α	*β*	*α*	*β*	*α*	*β*
Breaks rules*	.62	.00	.94	-.30	.36	.27
Doesn't seem to feel guilty after misbehaving	.53	.20	.93	-.35	.39	.25
Hangs around with others who get in trouble	.68	.30	.81	-.31	.39	.35
Lies or cheats	.51	.26	.91	-.01	.20	.80
Steals at home	.72	.63	-	-	.52	.96
Steals outside the home	.87	.70	-	-	.61	1.19
Steals	-	-	.78	.77	-	-
Swears	.47	.15	.78	.09	.22	-.29
Truant	.89	.67	.68	.21	.64	.71
Tardy to school or class	-	-	.66	-.07	-	-
	Wave 3					
	CBCL		**TRF**		**YSR**	
	α	*β*	*α*	*β*	*α*	*β*
Breaks rules*	.77	.00	-	-	.61	.19
Doesn't seem to feel guilty after misbehaving	.67	.17	-	-	.35	.53

Table S3.4 (cont'd)

Hangs around with others who get in trouble	.82	.35	-	-	.50	.33
Lies or cheats	.71	-.07	-	-	.20	.86
Steals at home	.94	.93	-	-	.49	1.28
Steals outside the home	.90	.91	-	-	.76	.89
Steals	-	-	-	-	-	-
Swears	.57	.11	-	-	.57	-.87
Truant	.77	.64	-	-	.55	.62
Tardy to school or class	-	-	-	-	-	-

Note: *The CBCL item "breaks rules" was used as the anchor item at each wave for model identification purposes.

Table S3.5. *Measurement invariance results*

	χ^2 difference test (df)	**RMSEA (90% CI)**	**CFI**	**TLI**
Configural Invariance	-	.012 (.010-.014)	.920	.915
Scalar Invariance	51.75† (28)	.012 (.010-.013)	.920	.916

Note: †Significant change in chi-square at p<.05.

Table S3.6. *Predictors of LCGA group membership for remaining pairings*

Conditional LCGA parameter estimates from 4-group model			
	Chronic v. escalating		
	Estimate	S.E.	*p*-value
ADI	-.007	.132	.955
Income	**-.475**	**.121**	**<.001**
Peer Delinquency	.311	.189	.101
Sex	**1.144**	**.266**	**<.001**
Race	.056	.278	.840
	Escalating v. desisting		
	Estimate	S.E.	*p*-value
ADI	.185	.141	.190
Income	.119	.123	.336
Peer Delinquency	**.902**	**.218**	**<.001**
Sex	**-.563**	**.255**	**.028**
Race	.034	.313	.913
	Desisting v. normative		
	Estimate	S.E.	*p*-value
ADI	**.275**	**.088**	**.002**
Income	**-.278**	**.093**	**.003**
Peer Delinquency	**.860**	**.154**	**<.001**
Sex	.286	.151	.058
Race	.237	.224	.291

Note: Multi-informant composite mean RB scores are the outcome measure. Bold font indicates $p < .05$.

Table S3.7. *LGM fit statistics and parameter estimates for caregiver reports of RB*

Unconditional LGM model fit statistics					
Model	-2lnL	χ^2 (df)	AIC	BIC	SABIC
Linear growth	**5101.84**	**-**	**5113.84**	**5147.62**	**5128.56**
Means model	5205.66	103.82† (3)	5211.66	5228.55	5219.02
Conditional LGM parameter estimates					
Parameter	Estimate	S.E.	p-value		
Intercept					
Mean	**.410**	**.049**	**<.001**		
Variance	**.225**	**.018**	**<.001**		
ADI	.029	.020	.141		
Income	**-.067**	**.021**	**.002**		
Peer Delinquency	**.134**	**.018**	**<.001**		
Sex	**.171**	**.034**	**<.001**		
Race	.071	.052	.167		
Slope					
Mean	-.011	.007	.150		
Variance	**.003**	**.000**	**<.001**		
ADI	-.001	.003	.868		
Income	.000	.004	.925		
Peer Delinquency	.006	.004	.102		
Sex	**-.021**	**.006**	**<.001**		
Race	.000	.008	.950		
Covariances					
Intercept with slope	**-.017**	**.002**	**<.001**		
Intercept 1 with Intercept 2					
MZ	**.245**	**.017**	**<.001**		
DZ	**.154**	**.018**	**<.001**		
Slope 1 with Slope 2					
MZ	**.003**	**.000**	**<.001**		
DZ	**.001**	**.000**	**<.001**		
Intercept 1 with Slope 2					
MZ	**-.021**	**.002**	**<.001**		
DZ	**-.012**	**.002**	**<.001**		
Residual variance	**.120**	**.010**	**<.001**		

<u>Note</u>: RB sum scores were computed based on maternal reports on eight CBCL items. Bold font indicates p<.05. †Significant change in chi-square at p<.05.

Table S3.8. *LCGA fit statistics and predictors of group membership for caregiver reports of RB*

Unconditional LCGA model fit statistics			
	AIC	BIC	Smallest group (% of sample)
# of groups			
1	5363.62	5380.50	100
2	4857.79	4891.54	24.9
3	4637.80	4688.43	14.5
4	4461.54	4529.04	3.3
5	4345.10	4429.48	2.5
6	4239.47	4340.73	2.3
Conditional LCGA parameter estimates from 4-group model			
Chronic v. normative			
	Estimate	S.E.	*p*-value
ADI	.166	.101	.102
Income	**-.400**	**.089**	**<.001**
Peer Delinquency	**.954**	**.118**	**<.001**
Sex	**.824**	**.221**	**<.001**
Race	.443	.238	.063
Escalating v. normative			
	Estimate	S.E.	*p*-value
ADI	.052	.186	.780
Income	-.210	.144	.145
Peer Delinquency	**1.020**	**.155**	**<.001**
Sex	-.220	.341	.519
Race	.515	.338	.128
Chronic v. desisting			
	Estimate	S.E.	*p*-value
ADI	-.018	.123	.884
Income	**-.289**	**.115**	**.012**
Peer Delinquency	**.650**	**.118**	**<.001**
Sex	.459	.254	.071
Race	.041	.290	.888

Note: RB sum scores were computed based on maternal reports on eight CBCL items. Bold font indicates $p<.05$.

Table S3.9. *Parameter estimates from conditional LGM with data in "long" format*

Parameter	Estimate	S.E.	p-value
	Intercept		
Mean	**.651**	**.061**	**<.001**
Variance	**.442**	**.042**	**<.001**
ADI	**.055**	**.025**	**.028**
Income	**-.159**	**.027**	**<.001**
Peer Delinquency	**.183**	**.026**	**<.001**
Sex	**.250**	**.046**	**<.001**
Race	.061	.063	.333
	Slope		
Mean	**.030**	**.010**	**.002**
Variance	**.005**	**.001**	**<.001**
ADI	.002	.004	.709
Income	**.012**	**.005**	**.007**
Peer Delinquency	**.021**	**.004**	**<.001**
Sex	**-.023**	**.007**	**.002**
Race	-.002	.010	.869
Intercept with slope	**-.037**	**.005**	**<.001**
Residual variance	**.326**	**.023**	**<.001**

Note: Multi-informant composite mean RB scores are the outcome measure. Bold font indicates $p<.05$.

Table S3.10. *LCGA fit statistics and predictors of RB group membership with 1 twin randomly selected per family (N = 1,030)*

Unconditional LCGA model fit statistics			
	AIC	BIC	Smallest group (% of sample)
# of groups			
1	3997.34	4012.15	100
2	3833.60	3863.22	39.4
3	3774.99	3819.41	5.0
4	3693.81	3753.03	8.1
5	3653.97	3728.00	4.2
6	3624.97	3713.81	2.8
Conditional LCGA parameter estimates from 4-group model			
	Chronic v. normative		
	Estimate	S.E.	*p*-value
ADI	**.573**	**.200**	**.004**
Income	**-.384**	**.163**	**.018**
Peer Delinquency	**1.939**	**.283**	**<.001**
Sex	.402	.317	.205
Race	.327	.378	.387
	Escalating v. normative		
	Estimate	S.E.	*p*-value
ADI	**.705**	**.207**	**.001**
Income	.109	.219	.618
Peer Delinquency	**1.671**	**.280**	**<.001**
Sex	-.291	.364	.423
Race	.351	.489	.474
	Chronic v. desisting		
	Estimate	S.E.	*p*-value
ADI	.211	.235	.369
Income	-.169	.166	.309
Peer Delinquency	**1.246**	**.262**	**<.001**
Sex	.102	.371	.784
Race	.066	.406	.870

Note: Multi-informant composite mean RB scores are the outcome measure. Bold font indicates $p<.05$.

Figures

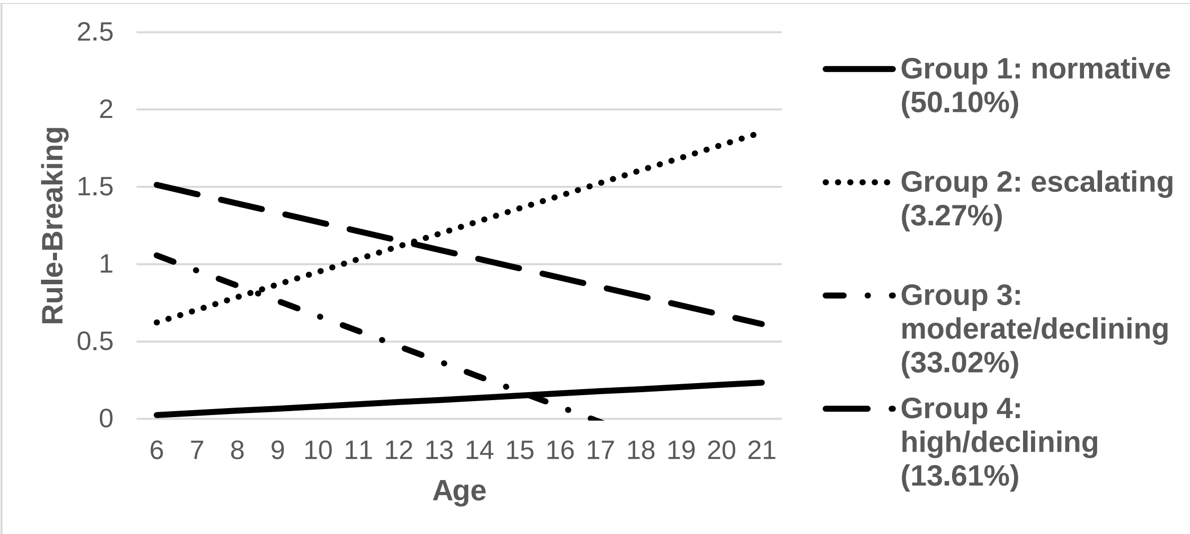


Figure S3.1. *Unconditional four-group model of RB trajectories based on maternal reports at all waves.*