DIFFERENTIAL SUSCEPTIBILITY TO PARENTING AND A FAMILY-CENTERED INTERVENTION IN EARLY CHILDHOOD: CONVERGENCE AND DIVERGENCE AMONG MULTIPLE METHODS AND MARKERS OF ASSESSING ENVIRONMENTAL SUSCEPTIBILITY

by

Chardée Ashley Galán

Bachelor of Arts in Psychology, University of California, Berkeley, 2012

Master of Science in Psychology, University of Pittsburgh, 2016

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Chardée Ashley Galán

It was defended on

August 9, 2019

and approved by

Daniel S. Shaw, Professor, Department of Psychology

Stephen B. Manuck, Professor, Department of Psychology

Jamie L. Hanson, Assistant Professor, Department of Psychology

Jay Belsky, Professor, Department of Human Ecology, University of California, Davis

Kathryn Lemery-Chalfant, Professor, Department of Psychology, Arizona State University

Dissertation Advisor: Daniel S. Shaw, Professor, Department of Psychology

DIFFERENTIAL SUSCEPTIBILITY TO PARENTING AND A FAMILY-CENTERED INTERVENTION IN EARLY CHILDHOOD: CONVERGENCE AND DIVERGENCE AMONG MULTIPLE METHODS AND MARKERS OF ASSESSING ENVIRONMENTAL SUSCEPTIBILITY

Chardée Ashley Galán, PhD

University of Pittsburgh, 2020

Although differential susceptibility theory has received increasing empirical support, several unresolved issues remain. The current study sought to address three particularly important questions concerning patterns of convergence or divergence across various markers and time scales of assessing differential susceptibility: (1) to what extent do different markers of susceptibility, specifically phenotypic (i.e., age 2 negative emotionality and effortful control) and genotypic characteristics, identify the same or different groups of children as being most susceptible to their caregiving environment during toddlerhood?; (2) are toddlers who respond with positive emotions to their parents' positive emotions the same children who also respond to their caregivers' negative display of emotions with negative emotions; (3) does differential reactivity to caregiving on a micro-time scale translate to differential susceptibility to positive and negative parenting across longer periods of development in relation to later conduct problems in middle childhood? Data were drawn from a randomized prevention trial conducted with youth and their primary caregivers followed prospectively from toddlerhood to middle childhood. Analyses using cumulative susceptibility scores showed that less than 1% of youth had elevated scores on all three markers of environmental susceptibility. As the majority of youth showed heightened susceptibility based only on one or two markers, findings provide initial evidence

suggesting divergence across different markers of environmental susceptibility. In regards to children's profiles of emotional reactivity during observed parent-child interactions, findings showed that a minority of children (5.2%) were differentially reactive "for better and for worse," responding with positive emotions to their parents' positive emotions and negative emotions to their parents' display of negative emotions. Finally, concerning the association between differential reactivity and differentially susceptibility, children who were differentially reactive to their parents' emotions during moment-to-moment interactions showed stronger longitudinal associations between negative and positive parenting in toddlerhood and later conduct problems. However, findings were consistent with diathesis stress, suggesting that short-term differential reactivity may not translate to differential susceptibility over many years. Collectively, findings underscore the importance of examining the extent to which different levels of analysis identify the same or different groups of children as being most susceptible to their environment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Conduct problems (CP) encompass aggressive, destructive, and rule-breaking behaviors that are moderately to highly stable from early childhood into adolescence and adulthood (Sitnick et al., 2017) and constitute the most common source of youth referrals to mental health clinics (Beauchaine et al., 2000; Reid, 1993). A wealth of longitudinal research has shown that relative to late starters who first engage in antisocial behavior in mid- to late-adolescence, early starters who display CP prior to age 10 engage in more serious forms of delinquency that persist across development (Moffitt et al., 2002; Moffitt & Caspi, 2001). The level of individual impairment and the significant costs to society associated with early-onset CP have motivated efforts to develop intervention and prevention programs that can be targeted in early childhood when child behavior and family relationships are more malleable (Reid et al., 2004).

To be appreciated is that the effects of risk factors and forces contributing to CP—such as parenting quality, peer relations, neighborhood safety and schooling experiences—may not be the same for all children, with some individuals being more and others less susceptible to the environmental regulation of their behavior. Thus, in addition to identifying proximal and distal contextual influences implicated in the development and maintenance of CP, it is also critical to identify children disproportionately affected by their developmental experiences and environmental exposures. Research on how individual characteristics modify environmental effects has traditionally been guided by the dual-risk (Sameroff, 1983) or diathesis-stress model (Monroe & Simons, 1991) of person X environment interactions. These frameworks posit that certain behavioral, physiological, or genetic characteristics increase vulnerability to environmental adversity (e.g., poverty, harsh punishment). However, other models of individual differences in

environmental sensitivity have recently been advanced, including vantage sensitivity (Manuck, 2011a; Sweitzer et al., 2013; Pluess & Belsky, 2013a), which is the converse of diathesis-stress (i.e., individuals vary in their susceptibility to supportive contextual conditions) and differential susceptibility. The latter posits that the same characteristics that make certain children disproportionately vulnerable to adversity may also increase susceptibility to supportive, enriching, or even just benign environmental conditions (Belsky, 1997a, 2005; Belsky, Bakermans-Kranenburg & van IJzendoorn, 2007; Belsky & Pluess, 2009; 2013a).

Studies of differential susceptibility can be distinguished based on three levels of analysis: (1) **marker of susceptibility**: genotypic markers versus endophenotypic markers versus phenotypic markers, (2) **method of assessing environmental susceptibility:** questionnaire data versus observational methods, and (3) **time scale:** long-term vs. short-term susceptibility to contextual influences. Although genotypic, endophenotypic, and phenotypic characteristics have individually garnered empirical support as markers of environmental susceptibility, albeit inconsistent, it is unclear whether they capture the same subgroups of children. Thus, the primary the aim of the current study is to examine whether phenotypic (i.e., negative emotionality and effortful control in toddlerhood) and genotypic characteristics converge or diverge in identifying youth who are most susceptible to their environment for "better and for worse."

In addition to the type of susceptibility marker, the methods used to assess susceptibility may also influence the ability to identify young children disproportionately affected by parenting and other environmental exposures for better and for worse. Evidence of differential susceptibility has been detected among studies assessing negative emotionality and effortful control using observational methods, as well as among studies incorporating questionnaire measures of these temperament dimensions. However, similar to the body of work on markers of susceptibility, it

remains unclear whether different methods of assessing environmental susceptibility are capturing the same subgroups of children. Thus, the current investigation seeks to examine whether parentreport and observational measures of temperament in toddlerhood converge or diverge in identifying youth who are most susceptible to positive and negative parenting in toddlerhood relation to later CP in middle childhood.

Finally, studies of differential susceptibility have primarily focused on long-term developmental changes in response to naturally occurring variation in the environment or interventions. However, a second, albeit less common, approach to studying differential susceptibility focuses on heightened susceptibility to the environment on a micro time scale, examining more immediate and short-term changes in behavioral functioning to a range of positive and negative stimuli. However, it remains unclear whether differential reactivity to minor environmental changes on a micro-time scale translates to differential susceptibility to one's environment across longer periods of development spanning months or years. By tracking moment-by-moment changes in children's emotional reactivity in response to their caregiver's display of positive and negative emotions during an observed interaction, the current study aims to evaluate whether short-term micro-level changes are consolidated into long-term susceptibility.

As the differential susceptibility perspective has important implications for understanding why interventions may be more or less effective for certain groups of children, research examining patterns of convergence and divergence across various markers, methods, and time scales is imperative. While knowledge of differential susceptibility could facilitate the design of personalized interventions, leading to a more optimal fit between individual characteristics and intervention type, this body of research also raises ethical concerns about discriminatory selective interventions based on children's susceptibility profiles. Examining the extent to which different

levels of analysis (e.g., different markers of susceptibility) identify the same or different groups of children as being most susceptible to their environment may help to guide screening procedures for interventions and educational services. For example, if findings from the present study indicate divergence across levels of analysis, then it will be important for screening procedures to incorporate data across multiple levels to ensure that children identified as susceptible at one level of analysis but not another are not mislabeled as "less susceptible" and denied services.

1.1 LITERATURE REVIEW

1.1.1 Level of Analysis: Markers of Differential Susceptibility

1.1.1.1 Phenotypic Markers of Susceptibility: Temperament

Temperament has been defined as "constitutionally based" individual differences in reactivity and self-regulation (Rothbart & Bates, 1998; p. 109). These differences appear to start early in life, are relatively stable across the life span, and are presumed to have a genetic basis (Goldsmith et al., 1987). While reactivity refers to the arousability of affective, motor, and sensory response systems, self-regulation involves the capacity to modulate such reactivity (Rothbart & Bates, 2006). Factor analyses of large datasets have yielded three broad temperament dimensions that have been consistently identified in studies of infants, children, and adults (Rothbart & Bates, 2006). The first two factors are related to emotional reactivity and describe positive/approach and negative/avoidance response tendencies. These factors broadly map onto the dimensions of positive emotionality and negative emotionality, respectively, and are typically measured in terms of an individual's threshold, intensity, and latency of affective arousal. A third temperament

dimension, often referred to as effortful control, is a regulative factor and refers to the capacity to inhibit a dominant response in favor of a subdominant one, enabling individuals to direct their attention and regulate their emotions and behaviors (Rothbart & Ahadi, 1994; Rothbart & Bates, 2006). Thus, effortful control can be thought of as a superordinate construct encompassing *attentional control* (i.e., the ability to focus and sustain attention as needed), *activational control* (i.e., the ability to perform an action when there is a strong desire to avoid it), and *inhibitory control* (i.e., the capacity to suppress an inappropriate behavioral or emotional responses).

Some of the earliest evidence of differential susceptibility to environmental influences emerged in research on temperament X parenting interactions (Belsky, 1997a). Early attempts to identify potential susceptibility markers called particular attention to negative emotionality and related attributes, including fear, distress, and difficult temperament (Belsky, 1997b, 2005; Belsky, Hsieh, & Crnic, 1998). Specifically, Belsky's (2005) review revealed that associations between early caregiving experiences and a variety of developmental outcomes were consistently greater for a subgroup of children characterized by a temperamental propensity for high negative emotionality. However, while the studies considered in this review were suggestive of differential susceptibility, they often reported statistical analyses that failed to differentiate disordinal interactions from interactions more consistent with diathesis stress or vantage sensitivity. As a result, it was unclear based on the research available at the time whether a for-better-*and*-for-worse parenting effect accounted for the greater variance in developmental outcomes explained by early caregiving experiences in children with higher levels of negative emotionality.

Since the publication of Belsky's (2005) review, research on differential susceptibility has blossomed, with an increasing number of studies providing empirical evidence that children high in negative emotionality are indeed differentially susceptibility to their environment consistent

with better-*and*-for-worse fashion (Belsky, Hsieh, & Crnic, 1998; Pitzer et al., 2011; Pluess & Belsky, 2010; Poehlmann et al., 2012). It is possible that negative emotionality may reflect a more sensitive nervous system from which experiences in the environment register especially strongly, irrespective of whether the experience is positive or negative (Aron & Aron, 1997; Aron, Aron, & Jagiellowicz, 2012; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & vanIjzendoorn, 2011). For example, in a sample of 109 preterm infants, Poehlmann et al. (2012) found that infant proneness to distress moderates the effects of maternal parenting at 9 months on externalizing behaviors at 36 months, such that infants observed to be more likely to become distressed exhibited more externalizing problems at 36 months relative to less distressed infants in the context of parenting characterized by high levels of criticism, anger, and frustration. Consistent with the differential susceptibility model, more easily distressed infants also exhibited fewer externalizing problems at 36 months than less distressed infants in the context of more positive parenting. The quality of parenting made no apparent difference for children scoring low in distress. Similarly, Bradley and Corwyn (2008) examined the interaction between infants' difficult temperament and early parenting in relation to teacher-rated externalizing problems in first grade using data from the NICHD Study of Early Child Care. The authors created a composite measure of parenting based on observed maternal sensitivity at 6, 15, 24, 36, and 54 months. Findings indicated that infants with difficult temperaments between 1 and 6 months were more affected by the quality of parenting they received than children with intermediate or low levels of difficult temperament (i.e., "average" and "easy" temperaments). Specifically, infants with difficult temperaments demonstrated more externalizing problems in first grade compared to infants with less difficult temperaments when exposed to less sensitive parenting, but fewer problems than infants with easy temperaments when they experienced more sensitive parenting.

It would be remiss, however, to not mention studies finding that infant negative emotionality operates in a manner consistent with diathesis stress rather than differential susceptibility (e.g., Kiff, Lengua, & Bush, 2011; Kochanska & Kim, 2013). Still, other studies find no interaction effects at all (e.g., Vitaro, Barker, Boivin, Brendgen, & Tremblay, 2006), while others find mixed results supporting both diathesis stress and differential susceptibility (Kochanska, Aksan, & Joy, 2007; Lengua, 2008). Thus, although the differential susceptibility hypothesis has been well supported in some studies using infant negative emotionality as the index of child susceptibility, others have not found results consistent with a differential susceptibility perspective.

Although it has received less empirical attention relative to negative emotionality, effortful control has also been implicated as a potential susceptibility marker (de Haan et al., 2010; Pitzer, Jennen-Steinmetz, Esser, Schmidt, & Laucht, 2011; Poehlmann et al., 2011). For example, Pitzer and colleagues (2011) found that compared to girls with average levels of effortful control in infancy and toddlerhood, those with low effortful control displayed more externalizing problems at age 8 if their mothers were observed to engage in less directive parenting (i.e., low levels of "restrictive guidance") when they were 2 years old. Consistent with differential susceptibility, girls with low effortful control also displayed less externalizing problems when exposed to higher levels of restrictive guidance. In contrast, the externalizing problems of girls with average effortful control were not moderated by differences in restrictive parenting (Pitzer et al., 2011).

Despite additional evidence supporting effortful control as a marker of contextual sensitivity (de Haan et al., 2010; Poehlmann et al., 2011), the collective body of research on this topic has yielded inconclusive results. Similar to negative emotionality, while some studies suggest that children *low* in effortful control respond more strongly to their environment (e.g., Pitzer et al.,

2011), others show children *high* in effortful control to be more sensitive (e.g., Halpern, Garcia Coll, Meyer, & Bendersky, 2001). Other studies report null effects, failing to show any evidence of an interaction between effortful control and parenting (Karreman, van Tuijl, van Aken, & Dekovic, 2009; Meunier, Roskam, & Browne, 2011; Olson, Sameroff, Kerr, Lopez, & Wellman, 2005).

Collectively, the discrepant findings in the aforementioned literature have led to both confusion and disagreement among the scientific community regarding the role of negative emotionality and effortful control as susceptibility factors, highlighting the need for a systematic synthesis of existing research on this topic. Thus, Slagt, Dubas, Deković, and van Aken (2016) used meta-analytic techniques to determine which model of environmental sensitivity (diathesis stress, differential susceptibility, vantage sensitivity) is best supported by current research on temperament x parenting interactions. The review incorporated 84 longitudinal studies that reported on an interaction between parenting and child temperament in a sample of youth 18 years or younger and focused specifically on negative emotionality, effortful control, surgency, or difficult temperament. Findings indicated that difficult temperament functioned as a marker of differential susceptibility, conferring increase sensitivity to both negative *and* positive parenting across a range of child outcomes, including externalizing and internalizing behavior and cognitive and social competence. However, when more specific facets of difficult temperament were examined, negative emotionality, but not surgency or effortful control, marked differences in susceptibility. Specifically, as expected, associations between negative parenting and negative child adjustment, as well as associations between positive parenting and positive child adjustment, were found to be stronger for children higher on negative emotionality. Further, while difficult temperament functioned as a marker of susceptibility regardless of when it was assessed,

differences in susceptibility as a function of children's negative emotionality were only present when this trait was measured during infancy. Additionally, findings revealed significant differences in the methodological quality of studies—while temperament-by-parenting interactions were detected for both questionnaires and observations, support for differential susceptibility was most evident when parenting was assessed using observational techniques.

The null findings for effortful control and surgency suggest that the heightened susceptibility conferred by difficult temperament is primarily driven by the "negative emotionality" component of this broad construct (Slagt et al., 2016). However, as noted by Belsky and Pluess (2009), negative emotionality has received significantly more empirical attention compared to other temperament dimensions. This disproportionate focus on negative emotionality may in part because of investigators' concerns with identifying characteristics that increase vulnerability to maladaptive outcomes as guided by the diathesis stress perspective. Indeed, while studies on negative emotionality contributed 31 effect sizes to Slagt et al. (2016)'s meta-analysis, there were only 13 effect sizes included in their review that represented an interaction between *effortful control* in infancy or toddlerhood and parenting. Thus, despite compelling evidence supporting negative emotionality as a susceptibility factor, it is premature to conclude that negative emotionality is the only or most important phenotypic marker of susceptibility. Relatedly, while Slagt et al. (2016)'s review suggests that effortful control is *not* a marker of contextual sensitivity, it is critical to remember that these findings represent the *current* state of knowledge. As these results may or may not hold as this body of research continues to grow, further investigation of other dimensions of temperament in addition to negative emotionality appears warranted. To this end, the present study examines both negative emotionality and effortful control as susceptibility factors.

1.1.1.2 Genetic Markers of Susceptibility

In addition to accounting for child-level characteristics, meta-analytic findings from twin and adoption studies suggest that CP are moderately to highly heritable (Rhee & Waldman, 2002). Prevailing biological theories focus on the dopaminergic system, which is associated with correlates of reward processing (Everitt & Robbins, 2000) and thought to underlie individual differences in the capacity to inhibit aggression and restrain impulses (Narvaes & Martins de Almeida, 2014). The canonical circuitry of reward processing, encompassing both striatal regions of the basal ganglia and prefrontal brain structures, is modulated by dopamine-releasing neurons projecting from the midbrain ventral tegmental area (Sweitzer et al., 2013). Heightened activation of the ventral striatum in response to reward-related stimuli, for instance, covaries with appetitive motivation (Beaver et al., 2006), impulsive disposition (Forbes et al., 2009), and steeper temporal discounting (Hariri et al., 2006). In turn, individual differences in the striatal response to reward may be predicted by polymorphic variation in genes regulating the synaptic availability of dopamine and postsynaptic receptor signaling (Forbes et al., 2009; Nikolova et al., 2011). Some of these same gene polymorphisms have also been variably implicated in risk for CP, delinquency and aggression (Guo et al., 2007; Chester, et al., 2016; Qadeer et al., 2017), and ADHD (Nikolaidis & Gray, 2010).

Important to appreciate is that even while carrying genetic variants is expected to increase risk of CP, many genetically susceptible youth demonstrate low levels of CP in early and middle childhood. As multiple pathways to resilient and maladaptive functioning are possible, it has been increasingly recognized that genetic variation often modifies sensitivity to environmental risk factors. Importantly, failure to adequately account for such gene-environment (GxE) interactions can mask effects of both genetic and environmental influences. Indeed, effects of *COMT, DRD2, DRD4*, and *DAT1* polymorphisms on CP have been shown to interact with various environmental

influences, such as maternal insensitivity, prenatal stress, and parent substance use (Kahn, Khoury, Nichols, & Lanphear, 2003; Propper, Willoughby, Halpern, Carbone, & Cox, 2007; Zohsel et al., 2014). For example, maternal insensitivity was associated with greater CP in toddlers, but only in the presence of the *DRD4* 7-repeat allele (Bakermans-Kranenburg & van IJzendoorn, 2006). Additionally, carriers of at least one *DRD4* 7-repeat allele demonstrated an increased risk for a lifetime diagnosis of Conduct Disorder/Oppositional Defiant Disorder in adolescence following exposure to higher levels of prenatal maternal stress. However, homozygous carriers of the *DRD4* 4-repeat allele proved insensitive to the effects of such environmental toxins (Zohsel et al., 2014).

While the majority of GxE studies of CP have been guided by the diathesis-stress model, implicitly or explicitly, evidence from individual GxE studies and meta-analysis suggest that certain dopamine-related genes may function in a manner consistent with differential susceptibility by moderating the association of both negative *and* positive environments with CP (Bakermans-Kranenburg & van IJzendoorn, 2011). For example, Bakermans-Kranenburg & Van IJzendoorn (2011) conducted a meta-analysis of 15 gene-environment studies that were published prior to May 2009 and involved dopamine-related genes in children up to 10 years old. Consistent with differential susceptibility, authors demonstrated that individuals with seemingly "vulnerable" dopamine-related genes are actually more susceptible to the environment "for better and for worse."

While Bakermans-Kranenburg and Van IJzendoorn's (2011) meta-analysis was theoretically innovative, it was limited by the availability of research on the "bright side" of GxE interactions at the time. Thus, Galán, Shaw, Belsky, and Manuck (under review) sought to extend this review, reconsidering it in the context of new empirical contributions that have added increasing complexity and specificity to the differential susceptibility literature over the past 7

years. Included studies (k = 58) reported on the interaction between an environmental context and dopaminergic polymorphisms (i.e., *DAT1* 40bp VNTR, *DRD4* 48bp VNTR, and/or *DRD2* Taq1) in relation to externalizing problems or positive development outcomes. Externalizing problems were operationalized to encompass a range of disruptive behaviors that are directed outwards towards one's environment, such as CP, aggression, oppositionality, and substance use. Findings indicated that contrary to expectations, individuals carrying putatively susceptible alleles were not at increased risk for externalizing problems or age-related correlates in the context of adverse environments, nor did they disproportionately benefit vis-à-vis externalizing problems in enriching and supportive environments. Thus, carriers of more susceptible and less susceptible alleles were equally affected by their environments, and there was no evidence of a significant GxE interaction, regardless of the racial composition of included studies or focus on adverse versus supportive environments.

In sum, the differential susceptibility hypothesis has received mixed support among studies examining genotypic markers of susceptibility. While some studies provide empirical support, other research has failed to corroborate such relationships, highlighting inconsistencies in the literature regarding the nature of GxE interplay with respect to dopaminergic genes and risk of CP. One perspective is that the null findings are consistent with growing empirical evidence casting doubt on the validity of prior GxE findings and the role of genotypic characteristics as susceptibility factors.

However, it is noteworthy that all meta-analytic reviews of differential susceptibility to date, and GxE interactions more broadly, have focused on single genetic polymorphisms that individually account for a small amount of variance in CP. Further, despite widespread practice, the use of single genetic variants contributes to multiple testing problems and ignores associations among SNPs in linkage disequilibrium. Additionally, by requiring an a priori understanding of the

biological pathways giving rise to a phenotype, this approach prioritizes functional polymorphisms, such as base-pair substitutions in gene coding regions that alter the amino acid sequence of a protein (Collier et al., 1996). Thus, Manuck and McCaffrey (2014) argue that a significant limitation of the candidate gene approach is that it "narrows the search space for genetic variation to components of prevailing biological models and, hence, rarely nominates more than a handful of the estimated 20,000 to 25,000 human genes…and neglect large expanses of the genome that do not code for protein" (p.44). Thus, although the overall effect sizes were non-significant in Galán et al. (under review), it is premature to dismiss genotypic characteristics as markers of environmental susceptibility based on the limitations of candidate gene studies.

In contrast to studies examining the influence of one genetic variant at a time, complex traits and behaviors such as CP are thought to have an underlying polygenic model of inheritance in which multiple genes account for differences in behavior (Risch & Merikangas, 1996; Wray et al., 2014). Hence, there are a growing number of studies incorporating polygenic scores formed from multiple SNPs as indices of genetic susceptibility. Polygenic scores are often formed based on a large number of SNPs drawn from a previous genome-wide association study (GWAS), with all SNPs below a certain significance threshold selected for inclusion. Accordingly, the current study capitalizes on advances in molecular genetic techniques by including a polygenic score that consists of GWAS-identified SNPs in *COMT, DRD2, DRD4*, and *DAT1* genes (Pappa et al., 2016). These scores are henceforth referred to as genome-wide polygenic scores (GPS).

While it has been less than a year since the current sample was genotyped, the lab has had initial success in forming meaningful polygenic scores. For example, Lemery-Chalfant and colleagues (2018) formed a differential susceptibility polygenic score based on SNPs that predicted depression/anxiety differences in identical twins in a GWAS (Lemery-Chalfant,

Clifford, Dishion, Shaw, & Wilson, 2018). This score moderated effects of a family-based intervention, the Family Check-Up (FCU), on children's reports of internalizing problems in a structured clinical interview. Specifically, significant differences between the FCU and control groups emerged at 0.5 SD above the mean on polygenic differential susceptibility, with the intervention leading to fewer internalizing symptoms compared to the control group for about 25% of the sample. Shaw, Galán, Lemery-Chalfant, Dishion, and Wilson (2018) also identified evidence of differential susceptibility, showing that in an ITT analysis comparing the effectiveness of the FCU in discriminating persistently high and persistently trajectories of conduct problems from ages 2 to 14, this intervention effect was qualified by an interaction involving genetic susceptibility (Shaw, Galán, Lemery-Chalfant, Dishion, & Wilson, 2018). Polygenic scores for aggression moderated FCU effects on trajectory class membership; those with greater polygenic scores were more likely to follow a persistently high trajectory unless they received the FCU, in which case they had a higher likelihood of being in the persistently low group. Conversely, those with lower polygenic scores did not differ in terms of trajectory group membership. Collectively, these studies underscore the viability of using polygenic scores in the current sample to example susceptibility to environmental influences in relation to meaningful behavioral outcomes.

Also important to consider is that although findings from Galán et al. (under review) yielded limited evidence for differential susceptibility based on *overall* effect sizes, significant heterogeneity in the distribution of effect sizes suggested that interactions between dopaminergic polymorphisms and the environment may be observed only under specific circumstances. Specifically, among studies focusing on *adverse* environments, GxE effects were identified when exposure to environmental adversity occurred prenatally but were absent at all other development periods from infancy/toddlerhood through late adolescence. Furthermore, relative to putatively

carriers of less susceptible alleles, children with more susceptible alleles demonstrated lower levels of externalizing problems and more positive developmental outcomes when exposed to salutary environments during infancy/toddlerhood. Such GxE effects were not present when exposure to supportive environments occurred during preschool, middle childhood, or adolescence.

Findings suggest that the prenatal period and infancy/toddlerhood may represent periods of increased susceptibility to environmental influences. Interestingly, results reported in Galán et al. (under review) parallel those of Slagt et al. (2016), the previously discussed meta-analysis that examined whether children vary in sensitivity to parenting based on individual differences in specific dimensions of temperament. The authors found that negative emotionality functioned as a marker of differential susceptibility, conferring increased sensitivity to both positive and negative parenting, but only when negative emotionality was assessed during infancy. This finding also was consistent with Belsky's (2005) narrative review. Thus, while Galán et al. (under review) and Slagt et al. (2016) evaluated different markers of susceptibility (i.e., genetic characteristics and temperament), their findings converge in highlighting the importance of the "early years of life, when biological systems are being laid down (e.g., Ganzel & Morris, 2011; Simpson et al., 2012)."

In their discussion of potential areas for future research, Slagt et al. (2016, p. 33) pose an important question, particularly relevant to this discussion: "…with meta-analyses now showing that susceptible children can be characterized by certain genotypes (Bakermans-Kranenburg & van IJzendoorn, 2011; van IJzendoorn et al., 2012), as well as by high difficult temperament/negative emotionality, the question that begs to be is whether researchers have been identifying the same children using these different susceptibility markers." Indeed, the parallel pattern of findings across meta-analyses suggest that genetic and temperament markers of susceptibility may in fact be identifying many of the same children. However, these speculations lack empirical support,

motivating the current study's aim to determine whether phenotypic (i.e., temperament) and genotypic characteristics are actually identifying the same "susceptible" individuals using different means.

1.1.2 Level of Analysis: Method of Assessing Differential Susceptibility

In addition to the type of susceptibility marker, the methods used to assess susceptibility may also influence the ability to identify young children disproportionately affected by parenting and other environmental exposures for better and for worse. At present, negative emotionality and effortful control in children are most commonly assessed through two broad measurement modalities, including caregiver-reported questionnaires and structured observations during laboratory-based tasks. Although both questionnaires and observational approaches offer several advantages to the study of child temperament, they are also accompanied by unique psychometric and practical challenges that are discussed below.

1.1.2.1 Level of Analysis: Questionnaire Measures of Temperament

Parent-reported questionnaires on dimensions of temperament, such as child negative emotionality and effortful control, ask caregivers to rate how their child is likely to react in a variety of situations. Parent report of temperament offers several advantages, as they are relatively inexpensive and quick to develop, administer, and analyze. Despite being subject to a range of biases, parent reports of child temperament might be more important than more objective indices if a parent's perception of their child is more impactful in influencing how they respond to their child's behavior. Parent ratings of child temperament also have the benefit of assessing a child's behavior across a number of contexts and over long spans of development.

Although the advantages of parent-reported measures of child temperament contribute to their popularity, they also have several limitations, including concerns of social desirability. Parent reports are also likely to be affected by personal biases related to parents' expectations, their attributions about their child, or their mood when reporting on their child's behavior (Monroe, 2008; Gardner, 2000; Prescott et al., 2000). The Infant Temperament Questionnaire (ITQ; Carey, 1973), for example, has been criticized for reflecting maternal characteristics such as anxiety and hostility more than infant temperament (Vaughn, Taraldson, Crichton, & Egeland, 1981). Relatedly, maternal characteristics have been found to explain more variance in maternal reports of temperament than observed infant behavior (Sameroff, Seifer, & Elias, 1982), suggesting that some mothers may be less accurate in reporting on their child's behavioral tendencies than others.

Parent report, as with all self-reported data, is also prone to biases in retrospective recall and is influenced by respondents' comprehension of instructions, questions, and rating scales. Differences in parents' reading levels, in turn, may result in different levels of cognitive load required to complete the survey. Indeed, increased subject burden has been found to result in poorer quality data because of lower response rates and greater errors (Bennett et al., 2003; Rolstad, Adler, & Rydén, 2011). Thus, differences in parent-reported effortful control, for instance, might not only reflect actual differences in children's ability to regulate their emotions and behaviors but also differences in the extent to which parents are affected by these biases when completing questionnaires. This may limit the validity of negative emotionality and effortful control as markers of environmental susceptibility, as parent reports of temperament may not perfectly correspond to a child's "true" disposition. However, as noted above, if such perceptions affect parenting practices, parent perceptions might be one of the most influential markers of differential susceptibility.

1.1.2.2 Level of Analysis: Observational Measures of Temperament

Observations of child temperament or parenting typically involve coding of micro-level processes, placing children (and sometimes parents) in multiple situations that vary in level of stress for the child to examine individual differences in child negative and positive emotionality, attention, and inhibitory control, among others. In some cases, the parent is removed from the room or the young child is asked to complete tasks with the examiner directing the assessment (e.g., Laboratory Temperament Assessment Battery (LABTAB). When parents are actively involved in the tasks, they are either asked to behave as they would normally toward the child (e.g., separating a toy from the child, working on a questionnaire while the child has nothing to do), or act in a constrained manner (e.g., still-face paradigm). Child behaviors and in some cases, parent behaviors (e.g., Martin high-chair task, Marvin cookie wait task), are subsequently coded.

Observational measures of temperament address many of the major weaknesses of questionnaire-based measures by capturing behavioral manifestations of negative emotionality and effortful control in a potentially less biased manner. Observational techniques are invaluable tools that allow researchers to directly observe and code objectively defined behaviors and emotions. Thus, child temperament and parent-child interactions can be coded such that the same operational definition of these constructs is applied to all families. This is in contrast to questionnaire data, which reflect a parent's subjective interpretation of the construct of interest (e.g., child effortful control).

Consistent with research examining child temperament as a marker of environmental susceptibility, genetic moderation of environmental context has also been found to be stronger in studies utilizing more objective measures to assess the environment relative to studies relying on self-reported questionnaires. For example, although recent meta-analyses found that the serotonin

transporter promoter polymorphism (*5-HTTLPR*) moderates the relationship between environmental adversity and depression (Karg, Burmeister, Shedden, & Sen, 2011; Uher & McGuffin, 2010), the method used to assess environmental adversity accounted for substantial heterogeneity in study findings. Specifically, an interaction between 5-HTTLPR and negative environmental conditions was most pronounced when environmental adversity was 'objectively' ascertained independently of participants' reports. Examples of such 'objective' measures of adversity included physical illness established by formal tests, well-known incidents of natural disasters, and exposure to child abuse confirmed by records from state child protection or social services agencies (Uher & McGuffin, 2010). In contrast, studies failing to replicate the GxE interaction more often relied on self-reported measures of adversity.

1.1.3 Level of Analysis: Short-Term vs. Long-Term Susceptibility

While studies employing parent-reported measures of temperament have significantly advanced our knowledge of differential susceptibility, research in this area has primarily focused on temperament and parenting at the group- rather than individual-level. However, the differential susceptibility model postulates that the *same* children that are most adversely affected by negative environments may also disproportionately benefit from positive environments. Thus, while the focus of differential susceptibility theory is on the nature of individuals, most studies in this area employ between-person designs, establishing relations at the level of the group or population. As it remains unclear whether relationships established at the group level are the same as those that operate at the level of processes occurring within individuals, within-person designs are needed in which the *same* individuals are exposed to both positive *and* negative environments. Thus, observational methods may be better suited for examining such within-person coupling compared

to questionnaire data, which more often restrict comparisons to *between* individuals. The current study seeks to advance research on differential susceptibility by including observational assessments of parent-child interactions in toddlerhood (in addition to parent reports of negative emotionality and effortful control) to determine whether the same children who respond with positive emotions to their parents' display of positive emotions are the same children who also respond to their caregivers' negative display of emotions with negative emotions.

Relatedly, studies of differential susceptibility have primarily focused on long-term developmental changes in response to naturally occurring variation in the environment or interventions. A second, albeit less common, approach to studying differential susceptibility focuses on heightened susceptibility to the environment on a micro time scale, examining more immediate and short-term changes in behavioral functioning to a range of positive and negative stimuli. Although this short-term approach tends to rely on experimental manipulations of the environment, it could also include immediate responses to naturally occurring events, such as a child's immediate affective reactions to their parent during moment-to-moment interactions (i.e., emotional reactivity). However, it remains unclear whether differential reactivity to minor environmental changes translates to differential susceptibility to one's environment across longer periods of development spanning months or years. By tracking moment-by-moment changes in children's emotional reactivity using observational techniques, it is possible to determine whether short-term micro-level changes are consolidated into long-term susceptibility. Thus, an additional aim of the present investigation is to examine whether the same children that react most strongly to their primary caregivers during observed moment-to-moment interactions are the same children who are also differentially susceptible to positive and negative parenting in toddlerhood predicting CP many years later in middle childhood.

1.1.4 Cumulative Susceptibility Scores

It is clear that questionnaire and observational measures of negative emotionality and effortful control both afford numerous strengths and weaknesses as assessment tools for identifying youth most susceptible to their environments. There are also many pros and cons to the use of genetic versus phenotypic characteristics for these purposes. However, while it is possible that a "gold standard" will be identified, it seems more likely that a combination of assessment techniques will be required. Thus, the current investigation examines the incremental validity of multi-informant and multi-method assessments, testing whether integrating information from a variety of sources incrementally contributes to the prediction of relevant criterion variables relative to scores taken from any one informant (i.e., parent-report vs. investigators' ratings) or method (i.e., questionnaire vs. observation; phenotypic vs. genotypic characteristics). Specifically, a cumulative score will be created for each child based on threshold cut points for observed emotional reactivity at age 2, parent-reported effortful control at age 2, and GPS. Children will receive a score of 0 if their score falls below the indicator's threshold and a score of 1 if their score is above this threshold. Scores on each of these indicators will be summed for each child, resulting in a single cumulative score that ranges from 0 to 3.

1.2 THE PRESENT STUDY

We sought to test the following aims and hypotheses using a low-income and racially diverse sample of youth who were initially recruited when they were 2 years old and have been assessed almost annually through age 16:

Aim 1: To examine whether phenotypic (i.e., age 2 negative emotionality and effortful control) and genotypic characteristics converge or diverge in identifying youth most susceptible to positive and negative parenting at age 3 for "better and for worse" in relation to CP in middle childhood.

Hypothesis 1: It is hypothesized that there will be moderate associations between observed negative emotionality at age 2 and parent-reported effortful control at age 2 as both are measurements of expressed dysregulated behavior. However, only modest associations are expected between children's genome-wide polygenic scores (GPS) and these two temperament dimensions, because genetic susceptibility may vary in its manifest expression in general and particularly by age 2. Further, using a person-oriented approach, it is expected that only a minority of children (i.e., $\leq$ 20%) will show higher levels of susceptibility on all three indicators of susceptibility.

Aim 2: To determine whether the same children that are more likely to respond with negative emotions to their parent's display of negative emotions are also the same children that are more likely to respond with positive emotions to their parent's display of positive emotions.

Hypothesis 2: It is hypothesized that there will be moderate associations between age 2 negative emotional reactivity and age 2 positive emotional reactivity. Further, using a person-oriented approach, it is expected that a minority of children will be differentially reactive (i.e., $\leq$ 20%), showing high levels of negative emotional reactivity *and* positive emotional reactivity.

Aim 3: To determine whether children differ in their susceptibility to negative *and* positive parenting in toddlerhood in relation to later CP based on their profile of phenotypic (i.e., negative emotionality and effortful control at age 2) and genotypic characteristics identified in Aim 1.

Hypothesis 3: It is hypothesized that children who demonstrate high levels of observed negative emotionality at age 2, high parent-reported effortful control at age 2, *and* high GPS will be most susceptible to positive and negative parenting at age 3 in relation to CP during middle childhood relative to children with lower-risk profiles. Specifically, it is hypothesized that children with more convergent profiles of susceptibility will show the highest CPs at age 10.5 when exposed to high levels of negative parenting, but the lowest CPs when exposed to high levels of positive parenting.

Aim 4: To determine whether children differ in their susceptibility to negative *and* positive parenting in toddlerhood in relation to CP in middle childhood based on their profile of emotional reactivity identified in Aim 2. Thus, this aim will examine the extent to which children's negative and positive emotional reactivity assessed at a micro, moment-to-moment scale translates into differential susceptibility to parenting and subsequent greater risk of early CP.

Hypothesis 4: It is hypothesized that children that react more strongly to their parents' emotions during moment-to-moment interactions will show stronger longitudinal associations between negative and positive parenting in toddlerhood in relation to CP at age 10.5. Specifically, compared to less reactive children, children high in emotional reactivity are expected to show higher levels of CP at age 10.5 when exposed to high levels of negative parenting but lower levels of CP when exposed to high levels of positive parenting.

2.0 METHOD

2.1 PARTICIPANTS

Participants for the current study include 731 caregiver-child dyads recruited between 2002 and 2003 from Women, Infants, and Children (WIC) Nutritional Supplement Clinics in the metropolitan areas of Pittsburgh, Pennsylvania, and Eugene, Oregon, and within and immediately outside of Charlottesville, Virginia (Dishion et al., 2008). Families with a child between the ages of 2 years, 0 months and 2 years, 11 months were invited to participate if they had family, child and/or socioeconomic risk factors for future behavior problems. To be deemed eligible for inclusion, families had to score above established clinical thresholds or at least one standard deviation above the normative mean in two of the three domains of risk: (a) familial risk (maternal depression – Center for Epidemiological Studies on Depression Scale; Radloff, 1977; or daily parenting challenges – Parenting Daily Hassles; Crnic & Greenberg, 1990; or self-report of substance or mental health diagnosis, or adolescent parent at birth of first child), (b) child risk (conduct problems – Eyberg Child Behavior Inventory; Robinson, Eyberg, & Ross, 1980; or high-conflict relationships with adults – Adult Child Relationship Scale; adapted from Pianta, 1995), and (c) socio-demographic risk (low education achievement – less than or equal to a mean of 2 years of post-high-school education between both parents and low family income using WIC criterion). In cases where two of the risk criteria were met on the basis of socio-demographic and family risk, then children were required to be above the standardized mean for CP or parent-child conflict to increase the probability that families would be motivated to modify child behavior.

Of the 1,666 families approached at WIC sites, 879 met the eligibility requirements (52% in Pittsburgh, 57% in Eugene, and 49% in Charlottesville), and 731 (83.2%) provided consent and assent. Of the 731 families who agreed to participate, 272 (37%) were recruited in Pittsburgh, 271 (37%) were recruited in Eugene, and 188 (26%) were recruited in Charlottesville. Children in the sample (49% female) were about 2.5-years-old (M = 29.9 months, SD = 3.2) at the time of the first assessment. Across sites, primary caregivers self-identified as belonging to the following ethnic groups: 50% European American, 28% African American, 13% biracial, and 9% other groups (e.g., American Indian, Native Hawaiian). Thirteen percent of the sample reported being Hispanic American. Over 96% of the primary caregivers at the initial assessment were biological mothers; thus, the terms "mother" and "primary caregiver" will be used interchangeably. During the initial screening period, more than two thirds of those families enrolled in the project had an annual income of less than $20,000, and the average number of family members per household was 4.5 (SD = 1.63). Forty-one percent of the sample had a high school diploma or general education diploma, and an additional 32% had 1-2 years of post-high school training.

2.1.1 Analytic Sample

Youth who were genotyped at age 14 comprise the subsample for the current study (n = 515 or 86.7% who participated in age 14 home visits at age 14). These adolescents were 50% female and belonged to the following racial/ethnic groups: 10% Latino, 30% African American, 48% European American, 5% Native American, 1% Asian American, and 6% other race or unknown race. With regards to study site, 129 were from Virginia, 184 were from Oregon, and 202 were from Pennsylvania.

Selective attrition analyses revealed no significant differences between those who did not versus did provide a saliva sample for genotyping with respect to parental education (high school diploma vs. no high school diploma), $\chi_2(1) = 0.40$, $p = .53$; minority racial status (Black vs. non-Black), $\chi_2(1) = 2.73$, $p = .10$; sex of child, $\chi_2(1) = 0.45$, $p = .50$; intervention status, $\chi_2(1) = 0.023$, $p = .88$; study site (Pittsburgh vs. non-Pittsburgh), $\chi_2(1) = 2.27$, $p = .13$, (Charlottesville vs. non-Charlottesville), $\chi_2(1) = 1.02$, $p = .31$; maternal depression at age 2, $t(590) = –0.003$, $p = .998$; externalizing problems at age 2, $t(591) = -1.204$, $p = 0.229$; or effortful control at age 2, $t(583) = -1.697$, $p = 0.090$.

2.2 PROCEDURES

Two- to three-hour assessments were conducted almost annually in families' homes with primary caregivers and their participating child from ages 2 to 10.5 years (missing only at child age 6), and again when children were 14 and 16 years. Each home assessment from ages 2 to 5 began by having the child engage in free play with age-appropriate toys while the mother completed a battery of questionnaires. After the free-play task (15 min), mother and child participated in a clean-up task (5 minutes), followed by a delay of gratification task (5 minutes), four teaching tasks (i.e., mother assisted the child put together a puzzle, build two towers, and play a board game; 3 minutes each), a second free-play (4 minutes) and clean-up task (4 minutes), the presentation of inhibition-inducing toys (2 minutes each), and a meal preparation/lunch task (i.e., primary caregiver prepared a meal for the child 20 minutes). All interactions were videotaped for later coding. When sufficiently old, target children also completed questionnaires regarding socio-

demographic characteristics, caregiver mental health, parenting, and their own behavior. The present study will utilize data collected at assessments occurring at ages 2, 3, 4, and 10.5.

The sample was derived from a larger randomized controlled trial investigating the prevention of children's CP, with 367 of 731 families (50.21%) randomly assigned to a family-based intervention, the Family Check-Up. To optimize study internal validity of the study, initial assessments at age 2 were completed prior to random assignment to the intervention (i.e., FCU) or control group (i.e., WIC care as usual). Families were then informed of their group status at the end of the assessment after examiners had completed global ratings of parenting and child functioning. In subsequent years, annual home assessments also took place prior to intervention for families assigned to the FCU, whereas control families participated only in annual assessments. Assessments were identical for control and intervention group participants. Parental written consent was obtained for all participants, and families were compensated for their time after each assessment. Institutional Review Board approval was received at each of the data collection sites.

Participants provided saliva samples with Oragene kits for genotyping. RUCDR Infinite Biologics at Rutgers University extracted and normalized the DNA. Samples were genotyped using the Affymetrix Axiom Biobank Array. 4,098,692 SNPs remained after basic post-imputation data cleaning. SNPs not in Hardy-Weinberg equilibrium ($p < 10^{-6}$), with a minor allele frequency less than 1%, and that fell within the Major Histocompatibility Complex on Chromosome 6 were removed. Copy-number variations were removed if they did not meet the 5% missing gene data threshold. 4,048,277 SNPs remained in imputed data after quality control procedures. Using the sliding window procedure in PLINK, we reduced linkage disequilibrium (LD) by screening out regions of long-range LD and local LD.

2.2.1 The FCU Intervention

The FCU is a brief (i.e., typically 3-5 sessions per year) family-centered intervention aimed at preventing early CP by enhancing parenting skills and addressing other domains that compromise parental functioning (e.g., parental psychopathology, social support; Dishion & Stormshak, 2007). Individually tailored to meet families' specific needs, the FCU consists of three, and in some cases, 4 components; 1) a get-to-know-you (GTKY) interview during which caregivers' concerns are explored, with a specific focus on family issues relevant to the child's behavior, 2) an ecological assessment involving videotaped parent-child interactions and questionnaires completed by caregivers, 3) a feedback session during which the parent consultant shares the results of the assessment with the caregiver(s), emphasizing family strengths while also using evidence-based motivational interviewing to promote positive parenting changes, and 4) follow-up treatment sessions which are offered if needed and if parents indicate that they are interested. The Everyday Parenting curriculum, a manualized intervention that focuses on parent management training, was used as a guide to follow-up interventions every year (Dishion et al., 2011). The FCU has extensive empirical support and is recognized as an evidence-based prevention program (e.g., SAMHSA 2015) for preventing youth behavioral problems and improving child and family adaptation from early childhood through adolescence (Dishion et al., 2008, 2014; Rubin et al., 2015; Shaw, Connell, Dishion, Wilson, & Gardner, 2009; Shaw et al., 2016).

Therapists in this randomized trial were found to have delivered the FCU with adequate fidelity, which was related to improvements in parenting and subsequent changes in children's problem behaviors between ages 2 and 4 (Smith et al., 2014). Engagement in the intervention was generally high as most families assigned to the intervention participated in the feedback sessions: 76.6% at age 2, 62.4% at age 3, 59.7% at age 4, 55% at age 5, 49% at age 7.5, 52% at age 8.5,

53.7% at age 9.5, and 45.2% at age 10.5. Across all years, families participated in an average of 4.54 (*SD* = 2.57) FCU feedback sessions; 24 families (6.5 %) never participated in a feedback; 69 (19.8 %) participated in one or two feedbacks; 63 families (17.2 %) participated in all eight FCU feedback opportunities; and, in total, 149 families (40.6 %) participated in 6 or more.

2.2.2 Measures

2.2.2.1 Polygenic Scores

Polygenic scores were computed based on a recent meta-GWAS of aggression in middle childhood (Pappa et al., 2016). Summary statistics including SNP reference number, risk allele, and *p*-value were drawn from this meta-GWAS. Polygenic scores were created using PRSice v2 (Euesden, Lewis, & O'Reilly, 2015) and PLINK v1.9 (Purcell et al., 2007) from overlapping SNPs in the meta-GWAS and those genotyped in the current sample and included 1200 SNPs at the $p <$ .01 threshold. In the current study SNPs were coded additively and unit-weighted such that greater values reflected greater predisposition for aggression.

2.2.2.2 Population Admixture

A Principal Components Analysis of all autosomal SNPs to represent population admixture using PLINK. We extracted the first 20 components, with the first component (PC1) having an eigenvalue of 28.84 and differentiating European-American and Latino groups from African-American groups, with most biracial participants falling in the middle. The second component (PC2) had an eigenvalue of 5.62 and differentiated non-Latino participants (European and African American) from Latino participants. The remaining components had eigenvalues ranging from 1.45 to 1.21 and were excluded from these analyses.

2.2.2.3 Emotional Reactivity

At child age 3, children and their primary caregivers participated in a series of videotaped parent-child interaction tasks. These interactions were subsequently coded using the Relationship Affect Coding System (RACS; Peterson, Winter, Jabson, & Dishion, 2008). The RACS is a micro-social coding system that reflects three dimensions of behavior (verbal, physical, and affect) for each of family member simultaneously. Verbal codes comprise positive, neutral, and negative talk and include verbal behavior change codes, such as positive structuring, neutral, and negative directives. Physical behaviors (e.g., handing each other objects) are coded as positive, neutral, and negative. Affect codes include anger, disgust, distress, ignoring, and positive affect. The "off" codes of no talk, no physical, and neutral affect are used when verbal, physical behavior, or affect streams are not observed. The RACS coding was carried out using Noldus Observer XT, Version 11.0 (Noldus Information Technology, 2012), which enables continuous coding of an interaction as the behaviors are observed. As such, the exact durations and frequencies of behaviors are captured.

We prioritized the use of affect codes (i.e., not verbal codes or physical codes) based on the study's focus on examining children's emotional reactivity to their parent's display of positive and negative emotions. To capture broad dimensions of negative and positive affect, anger, disgust, distress, and ignoring were categorized as negative emotions, while validation and positive affect were categorized as positive emotions. Further, to examine emotional reactivity, changes in emotions were examined between five-second intervals and will range from -2 to 2:

1) Score of -2: 'When one dyad member (i.e., parent or child) becomes more positive, the other member becomes more negative'

2) Score of -1: ‘When one dyad member becomes more positive, the other remains stable’ or ‘When one dyad member remains stable, the other becomes more negative’
3) Score of 0: ‘When one dyad member remains stable, the other also remains stable’, ‘When one dyad member becomes more positive, the other also becomes more positive’, or ‘When one dyad member becomes more negative, the other also becomes more negative’
4) Score of 1: ‘When one dyad member remains stable, the other becomes more positive’, ‘When one dyad member becomes more negative, the other remains stable’
5) Score of 2: ‘When one dyad member becomes more negative, the other becomes more positive’

Thus, both parents and children received a code for each five-second segment of an interaction. Categories were collapsed (e.g., -2 and -1 into -1 and 1 and 2 into 1) as closer inspection of the data revealed an insufficient number of observations within categories.

A state space grid method was used to summarize the dynamic exchanges between the caregiver and child – i.e., the relationship between a parent’s display of positive and negative emotions and their child’s emotional expressions on a moment-to-moment timescale (i.e., ever 5 seconds). In the state space grid, the parent’s coded behavior was selected as time T, and the child’s reaction as T+1. This data structure required forming a set of priority rules from the three parallel streams. For instance, if the caregiver smiled and at the same time was observed to be saying something negative to the child, the negative verbal code trumped the smile, a positive affect code. The end result allows us to analyze the trajectory of sequential dyadic exchanges that have been

graphed on the state space grid (Hollenstein, 2007). State space grids have previously been adapted for the study of caregiver-child interactions and their relationship with child psychopathology (e.g., Granic, Hollenstein, Dishion, & Patterson, 2003; Hollenstein, Granic, Stoolmiller, & Snyder, 2004).

Trajectories of dyadic interaction patterns across time were computed for each mother-child dyad using the software program GridWare 1.15a (Lamey, Hollenstein, Lewis, & Granic, 2004). These trajectories were used to calculate the following variables the percentage of interaction time that children changed their emotions *given* prior changes in their mothers' emotions.

2.2.2.4 Negative Parenting

Negative parenting at age 3 was assessed using a composite index (see Moilanen et al., 2009) formed from three duration proportions from the Relationship Process Code (RPC; Jabson, Dishion, Gardner, & Burton, 2004) and five items from the Coder Impressions Inventory (COIMP; Dishion, Hogansen, Winter, & Jabson, 2004). First, a team of research assistants coded videotaped family interaction tasks using the RPC, a third-generation code derived from the Family Process Code (Dishion et al., 1983), which has been used extensively in prior research. The RPC includes coding of both verbal and behavioral displays, which are categorized as being positive, negative, or neutral. Three RPC codes were aggregated to form an observed negative parenting construct: the duration proportions of parental negative verbal, negative directive, and negative physical behavior. Fifteen percent of videotapes were coded twice, and the average team percent agreement was .87, $\kappa = .86$. Then, following *micro-social* coding, coders completed *macro-social* ratings of videos using the COIMP. Harsh parenting was assessed by COIMP items that assessed parents' provision of developmentally-inappropriate reasons for children's behavior change, displays of

anger or annoyance with the child, criticizing or blaming the child for family problems, use of physical discipline, ignoring/rejecting the child, and messages about the child's worthlessness. The three RPC codes and five COIMP macro ratings were standardized and summed in order to create a composite index of negative parenting (Cronbach's $\alpha = .75$).

2.2.2.5 Positive Parenting

Positive parenting at age 3 was assessed using a composite index (see Dishion et al., 2008) formed from the RPC, COIMP, and Home Observation for Measurement of the Environment inventory (HOME; Bradley, Corwyn, McAdoo, & Garcia-Coll, 2001). In detail, the following items were entered into the positive behavior support scores:

1) Parent Involvement. This measure is based on the home visitor's rating of the primary caregiver's involvement, which used three items from the HOME assessing whether the parent looks at the child, talks to the child, and/or structures the child's play periods.
2) Positive Reinforcement. This measure is based on videotape coding (durations) of caregivers prompting and reinforcing young children's positive behavior as captured in the following RPC codes: positive reinforcement (verbal and physical), prompts and suggestions of positive activities, and positive structure (e.g., providing choices in a request for behavior change).
3) Engaged Parent – Child Interaction Time. This measure reflects the average length of parent-child sequences that involve talking or physical interactions such as turn taking or playing a game. Thus, the average duration of episodes that included consecutive parent – child exchanges involving RPC codes such as Talk and Neutral Physical Contact were used to define these episodes.

4) Proactive Parenting. Using six items from the COIMP (α= .835), videotape coders rated each parent on his or her tendency to anticipate potential problems and to provide prompts or other structural changes to avoid young children becoming upset and/or involved in problem behaviors. Sample items included, "*Parent redirects the child to more appropriate behavior if the child is off task or misbehaves*" and "*Parent uses verbal structuring to make the task manageable*" (Cronbach's α at age 2 = .835, Cronbach's α at age 3 = .873).

2.2.2.6 Effortful Control

The 13-item effortful control (EC) subscale of the Child Behavior Questionnaire (CBQ; Rothbart, Ahadi, Hershey, & Fisher, 2001) was used to assess behavioral self-regulation at age 2. Items address self-regulatory capacities such as delaying an impulse, modulating activity level, and following directions (e.g., "*Is good at following instructions*" and *"can easily stop an activity when s/he is told 'no.'*"). The questionnaire's EC subscale is often consolidated in measures of effortful control and is widely used in developmental research on temperament (Ahadi et al., 1993; Olson et al., 2005). Mothers responded to each item on a 7-point scale, ranging from 0 = *extremely untrue of child* to 6 = *extremely true of child*. Scale scores were computed by averaging all numeric responses. Mothers could also indicate whether any items were not applicable to their child, and these items were treated as missing (i.e., were not averaged into scale scores). Scale scores were not computed if data were missing for three or more items. Maternal reports of EC were significantly, albeit modestly, correlated with observed EC during a delay of gratification task in which the child was asked to wait for a cookie at 3 years ($r = 0.21$, $p<0.01$).

2.2.2.7 Conduct Problems

Adopting a multi-method, multi-informant approach, three aspects of CP will be examined: (1) parent-report of CP, (2) youth-report of CP, and (3) teacher-report of CP. Primary caregivers rated children's CP at age 10.5 using the 113-item Child Behavior Checklist (CBCL) for Ages 6–18 (Achenbach & Rescorla, 2001). Responses were scored on a 3-point-response scale (0 = not true, 1 = somewhat or sometimes true, 2 = very true or often true). The CBCL was normed on a large sample of children ages 6–18 and possesses excellent test-retest and interrater reliability, as well as adequate to excellent internal consistency (Achenbach & Rescorla, 2001). The total raw score from the 13-item CP DSM-IV oriented scale was used for analyses. Sample items from this scale include "destroys property belonging to others" and "threatens people."

At age 10.5, youth rated the frequency of their CP using the Self-Report of Delinquency Questionnaire (SRD; Elliott, Huizinga, & Ageton, 1985). Using a 3-point scale (1 = never, 2 = once/twice, 3 = more often), children rated the extent to which they had engaged in aggressive and deviant behaviors (e.g. stealing, skipping school without an excuse) in the past year, with sample items including "In the past year, have you on purpose broken or damaged or destroyed something belonging to your parent or other people in your family?" and "In the past year, have you hit other students or got into a physical fight with them?" Due to the age of children, items regarding sexual risk taking were removed. Additionally, substance use was assessed in a separate measure, and as such, related items were removed from the SRD for parsimony. Items were summed to create an index of participants' self-reported conduct problems ($\alpha = .67$).

Teacher-report of youth CP at age 10.5 was assessed with the Teacher Report Form (TRF; Achenbach and Rescorla, 2001), a 112-item rating scale that is parallel to the CBCL. Teachers rated each item on a 3-point scale (0 = not true, 1 = somewhat or sometimes true, 2 = very true or

often true), and the total raw score from the 13-item CP DSM-IV oriented scale was used for analyses.

2.2.2.8 Covariates

Child sex (0 = *female*; 1 = *male*), intervention status (0 = *control*; 1 = *intervention*), and age in months were included as covariates in the present study. Additionally, geographic location was accounted for using two dummy variables comparing Eugene, Oregon, and Charlottesville, Virginia, to Pittsburgh, Pennsylvania. Additional covariates included the family monthly income, and the first two ancestry principal components, PC1 and PC2.

2.3 DATA ANALYTIC PLAN

Aim 1: Bivariate correlations were computed to test the hypothesis that there are moderate associations between observed negative emotionality at age 2 and parent-reported effortful control at age 2. Bivariate correlations were also computed to test the hypothesis that there are modest associations between children's genome-wide polygenic scores (GPS) and these two temperament dimensions. Additionally, more exploratory analyses examined the incremental validity of multi-informant and multi-method assessments. Specifically, a cumulative score was created for each child based on threshold cut points for observed negative emotionality at age 2, parent-reported effortful control at age 2, and GPS.

To test the hypothesis that a minority of children show high levels of susceptibility based on phenotypic *and* genotypic characteristics, latent profile analysis (LPA; Muthen & Muthen, 2000) was implemented using *MPlus 7.4* (Muthen & Muthen, 2012), with full information maximum likelihood employed to accommodate missing data. Indicators for the LPA included

observed negative emotionality at age 2, parent-reported effortful control age 2, and GPS, and LPA facilitated the identification of subsets or profiles of children who demonstrate similar patterns of susceptibility based on these indicators. The following criteria were considered when selecting the optimal number of profiles in combination with parsimony and theoretical consistency:

1) The Bayesian Information Criterion (BIC), the sample-sized adjusted version of the BIC (Adj BIC), and the Akaike Information Criterion (AIC), with smaller values indicating better model fit
2) Entropy, which reflects the effective classification of individuals into their most likely profile. Entropy values range from 0 to 1, with values closer to 1 indicating greater confidence in appropriate classification (Celeux & Soromenho, 1996), and
3) The Bootstrapped Likelihood-Ratio Test (BLRT), which provides a statistical comparison of the fit of a given model with a model of one fewer profiles. Additional criteria for selecting the optimal number of profiles included high posterior probabilities close to 1 for profile membership (Andruff et al., 2009; Jung & Wickrama, 2008).

Aim 2: Bivariate correlations were computed to test the hypothesis that there are moderate associations between child negative emotional reactivity at age 2 and positive emotional reactivity at age 2. Next, to test the hypothesis that a minority of children are differentially reactive to their parent's display of positive and negative emotions at age 2, LPA was implemented to identify subsets of children who show similar profiles of emotional reactivity (e.g., high on positive *and* negative emotional reactivity; low on positive emotional reactivity and high on negative emotional reactivity). The negative emotional reactivity and positive emotional reactivity variables

previously described in the methods section were used as profile indicators. The optimal number of profiles was determined based on the criteria outlined in the data analytic section for Aim 1.

Aim 3: To determine whether children differ in their susceptibility to negative *and* positive parenting in toddlerhood in relation to later CP based on their profile of phenotypic (i.e., negative emotionality and effortful control at age 2) and genotypic characteristics, the 3-step approach (Vermunt et al., 2010) was implemented in MPlus:

1) Step 1: Identifying the optimal number of profiles
2) Step 2: Saving children's predicted profile membership ("susceptibility profile") from Aim 1
3) Step 3: Computing regression analyses to investigate the main effect of susceptibility profiles, the main effect of parenting at age 3, and the interaction between children's susceptibility profile and parenting in relation to CP at age 10.5.

As positive parenting and negative parenting were tested in independent models, interactions with susceptibility profiles were calculated separately for each parenting dimension. Analyses accounted for the main effects of all covariates, including child sex, age in months, ancestry principal components, study site, and family monthly income. As recommended by Keller (2014), initial models also included two-way interactions between covariates and susceptibility profiles. Main effects of all covariates were retained in final models regardless of significance, but two-way interactions between polygenic score and covariates were trimmed if they were nonsignificant. As outlined by the 3-step approach, the final step included manually fixing logit values for misclassification error rates from step #1 when computing regressions. By accounting for the degree of uncertainty in classification, the three-step approach ensures that class membership is not affected by the inclusion of covariates, in contrast to the one-step approach that

can result in substantial changes to the identified profiles once covariates are introduced into the model. Analyses for Aim 3 were first computed for the entire sample and then followed by analyses computed separately for Caucasian and non-Caucasian youth. While the latter analyses included African American and Latino youth together in an effort to maximize power, models were also computed for African American youth only as they comprised over half of the "non-Caucasian" group.

Aim 4: To examine the extent to which children's negative and positive emotional reactivity assessed at a micro, moment-to-moment scale translates into differential susceptibility to parenting and subsequent greater risk of CP in middle childhood, analytic models outlined in Aim 3 were recomputed replacing children's "susceptibility profile" with their profile of "emotional reactivity" identified in Aim 2. Specifically, the 3-step approach was implemented, with regression models testing the interaction between children's emotional reactivity profiles at age 2 and their exposure to positive and negative parenting at age 3 in relation to CP at age 10.5. These models again accounted for the main effects of emotional reactivity, the main effect of parenting, the main effect of covariates, and two-way interactions between covariates and susceptibility profiles. Full information maximum likelihood estimation (FIML) was used to handle missing data.

3.0 RESULTS

3.1 AIM 1

3.1.1 Descriptive Statistics and Zero-Order Correlations

Descriptive statistics and bivariate correlations for all study variables are presented in Tables 1 and 2, respectively. Consistent with Hypothesis 1, observed negative emotionality (NE) at age 2 was inversely related to parent-reported effortful control (EC) at age 2. Unexpectedly, neither age 2 EC nor age 2 NE were related to aggression genome-wide polygenic scores (GPS), teacher-reported conduct problems (CP) at age 10.5, or youth-reported CP at age 10.5. Age 2 EC but not age 2 NE was significantly associated with parent-reported CP at age 10.5; youth with higher levels of parent-reported EC as toddlers were rated by their parents as having fewer CP in middle childhood. Higher levels of observed positive parenting at age 3 were related to lower levels of parent- and teacher-reported CP at age 10.5 but not related to youth-reported CP at age 10.5. In contrast, negative observed parenting at age 3 predicted greater teacher- and youth-reported CP at age 10.5 but was unrelated to parent-reported CP at age 10.5. Girls showed higher levels of parent-reported EC and lower levels of observed NE than did boys at child age 2, as well as lower levels of parent- and teacher-reported CP at age 10.5. There were no gender differences with respect to youth-reported CP in middle childhood.

Table 1. Descriptive statistics for analytic sample (*N* = 515)

	N	*Minimum*	*Maximum*	*Mean*	*SD*
Child Effortful Control (Age 2; Parent-Report)	508	1.440	6.150	3.990	0.793
Child Negative Emotionality (Age 2; Observed)	513	0.000	4.785	0.509	0.923
Positive Parenting (Age 3; Observed)	471	2.910	8.730	5.896	0.999
Negative Parenting (Age 3; Observed)	471	1.000	4.750	1.800	0.707
Polygenic Score (Differential Susceptibility; p <.001)	515	0.372	0.617	0.493	0.043
Polygenic Score (Middle Childhood Aggression; p <.001)	515	0.412	0.511	0.463	0.017
Ancestry PC1	515	-0.039	0.081	-0.000	0.044
Ancestry PC2	515	-0.245	0.027	0.000	0.044
Conduct Problems (Age 10; Parent-Report)[a]	463	50.000	90.000	57.960	8.709
Conduct Problems (Age 10; Teacher-Report)[a]	301	50.000	86.000	56.910	7.930
Conduct Problems (Age 10; Youth-Report)	424	0.000	25.000	1.790	2.504
Study Site	515	Pittsburgh, PA: *n* = 202 (39.2%) Charlottesville, VA: *n* = 129 (25%) Eugene, OR: *n* = 294 (57.09%)			
Family Income (Gross Yearly)	515	Less than $20,000: *n* = 345 (66.99%) $20,000-$39,000: *n* = 170 (33.01%)			
Intervention	515	Control: *n* = 256 (49.7%) Family Check-Up Intervention: *n* = 259 (50.3%)			
Child Sex	515	Female: *n* = 253 (49.1%) Male: *n* = 262 (50.9%)			
Child Race	515	Non-Hispanic Caucasian: *n* = 232 (45%) African American or Biracial: *n* = 213 (41.4%) Latino: *n* = 60 (11.7%) Other: *n* = 10 (1.9%)			

***Note*. Ancestry PC1 and PC2 are principle components accounting for genetic variation due to race/ethnicity.**

[a] Although raw CBCL were used for models to avoid potential age and gender corrections, for ease of interpretation, we present t- scores on the CBCL measures.

Table 2. Zero-order correlations

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. Intervention	1.0														
2. Child Sex	-.037	1.0													
3. Effortful Control (2 yrs.)	-.056	-.110*	1.0												
4. Negative Emotionality (2 yrs.)	.028	.094*	-.117*	1.0											
5. Negative Emotional Reactivity (2 yrs.)	-.007	-.040	.057	.077†	1.0										
6. Positive Emotional Reactivity (2 yrs.)	.020	-.059	.043	.025	.587*	1.0									
7. Pos. Parenting (3 yrs.)	.087†	.024	.072	.046	.050	.100*	1.0								
8. Neg. Parenting (3 yrs.)	-.027	.026	-.077†	.034	-.079†	-.101*	-.501*	1.0							
9. DS Polygenic Score	.028	-.045	-.047	-.053	-.049	-.007	-.035	.039	1.0						
10. Aggression Polygenic Score	.002	-.043	.059	-.028	-.042	-.050	-.003	-.015	-.028	1.0					
11. Ancestry PC1	-.028	-.012	.074†	-.002	-.158*	-.150*	-.168*	.098*	.226*	-.152*	1.0				
12. Ancestry PC2	.000	.039	-.056	.003	-.110*	-.040	-.152*	.074	-.017	.071	.000	1.0			
13. Conduct Problems (10.5 yrs; Parent-Report)	.059	.151*	-.167*	.033	-.023	-.040	-.116*	.091	-.053	-.102*	.051	.075	1.0		
14. Conduct Problems (10.5 yrs; Teacher-Report)	-.036	.234*	-0.30	.100†	.024	.003	-.138*	.208*	-.087	-.001	.214*	.057	.311*	1.0	
15. Conduct Problems (10.5 yrs; Youth-Report)	.019	.077	.057	.008	-.048	-.049	-.058	.105*	-.067	-.068	.143*	.018	.221*	.221*	1.0

Note. **Intervention: 0 = Control, 1 = Family Check-Up Intervention; Child Sex: 0 = Female; 1 = Male; Ancestry PC1 and 2 are principle components accounting for genetic variation due to race/ethnicity.**

†$p < .10$; * $p \leq .05$

3.1.2 Step 1: Identifying Latent Profiles of Susceptibility Based on Temperament and Genotype

After examining bivariate correlations, latent profile analysis (LPA) was implemented to identify subsets of children who demonstrate similar patterns of susceptibility based on observed negative emotionality at age 2, parent-reported effortful control at age 2, and aggression genome-wide polygenic scores (GPS). Supplementary Table 1 (Table S1) reports fit indices for 1-4 profile models, with higher entropy and lower BIC, AIC, and loglikelihood values suggesting an improvement in model fit when estimating *k* versus *k*-1 profiles. The four-profile model yielded the lowest BIC, adjusted BIC, AIC, and log likelihood values, and although the three-profile model had the highest entropy, both the three- and four-profile models yielded groups consisting of less than 5% of the sample — 4.30% and 4.17%, respectively. Based on these aforementioned inconsistencies across fit indices, plots were carefully reviewed for conceptual interpretability and clarity. As shown in Figures S1 and S2, regardless of the number of profiles estimated, GPS were nearly identical and importantly, limited in variability across all profiles. As limited variability in GPS (*SD* = 0.017; variance = .0003; Table 1) precludes the identification of meaningful subgroups on the basis of this indicator, using LPA with all of the planned variables was therefore deemed untenable. Thus, a LPA was re-computed using only observed NE at age 2 and parent-reported EC at age 2, and the aggression GPS was included in subsequent moderation analyses as a continuous variable. As shown in Table 3, AIC, BIC, adjusted BIC, BLRT, and entropy values supported a three-profile solution using NE and EC at age 2. However, despite improved model fit when estimating three versus two profiles, the three-profile model resulted in a group that consisted of

4.3% of participants. Visual inspection of the two-profile solution indicated that this model included an Average EC/Low NE profile and an Average EC/Moderate NE profile (Figure 2). The three-profile solution included an additional profile of children with Average EC/High NE (Figure 1). Despite the relatively small size of the Average EC/High NE profile (4.3% of sample), it appeared to be distinct from the other two other profiles. Further, as children high in NE are conceptually meaningful for the purposes of the current study, the three-profile solution was retained for further analyses.

Table 3. Aim 1: Fit indices for one to three group latent profile models based on effortful control and negative emotionality at Age 2

# Profiles	Loglikelihood	AIC	BIC	Adjusted BIC	Entropy	*p-value* BLRT	Class Sizes (%)
1	-616.186	1236.371	1244.634	1238.286	N/A[a]	N/A[a]	100
2	-440.072	888.143	904.668	891.973	0.973	p <.001	90.03/9.97
3	**-344.423**	**700.847**	**725.634**	**706.592**	**0.983**	***p* <.001**	**86.74/8.93/4.33**
4	-836.389	1698.778	1752.874	1711.614	0.946	p = .296	6.96/13.08/4.00/75.95

***Note*. AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; BLRT = Bootstrap Likelihood Ratio Test for k versus k-1 classes; Bolded text indicates best fitting model chosen.**

[a] Entropy and BLRT are not available for one-class models.

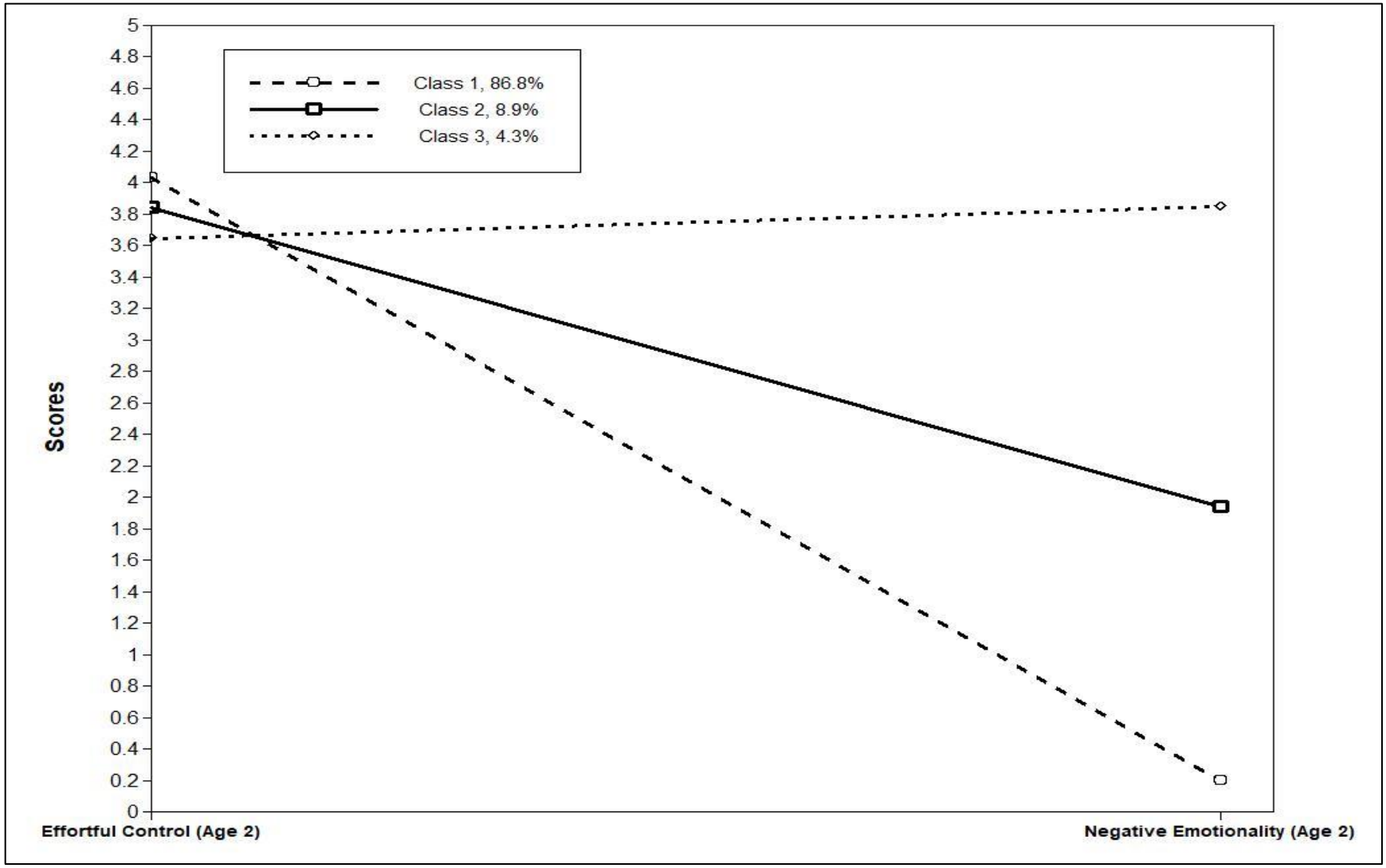

Figure 1. Estimating three latent profiles based on effortful control and negative emotionality at age 2.

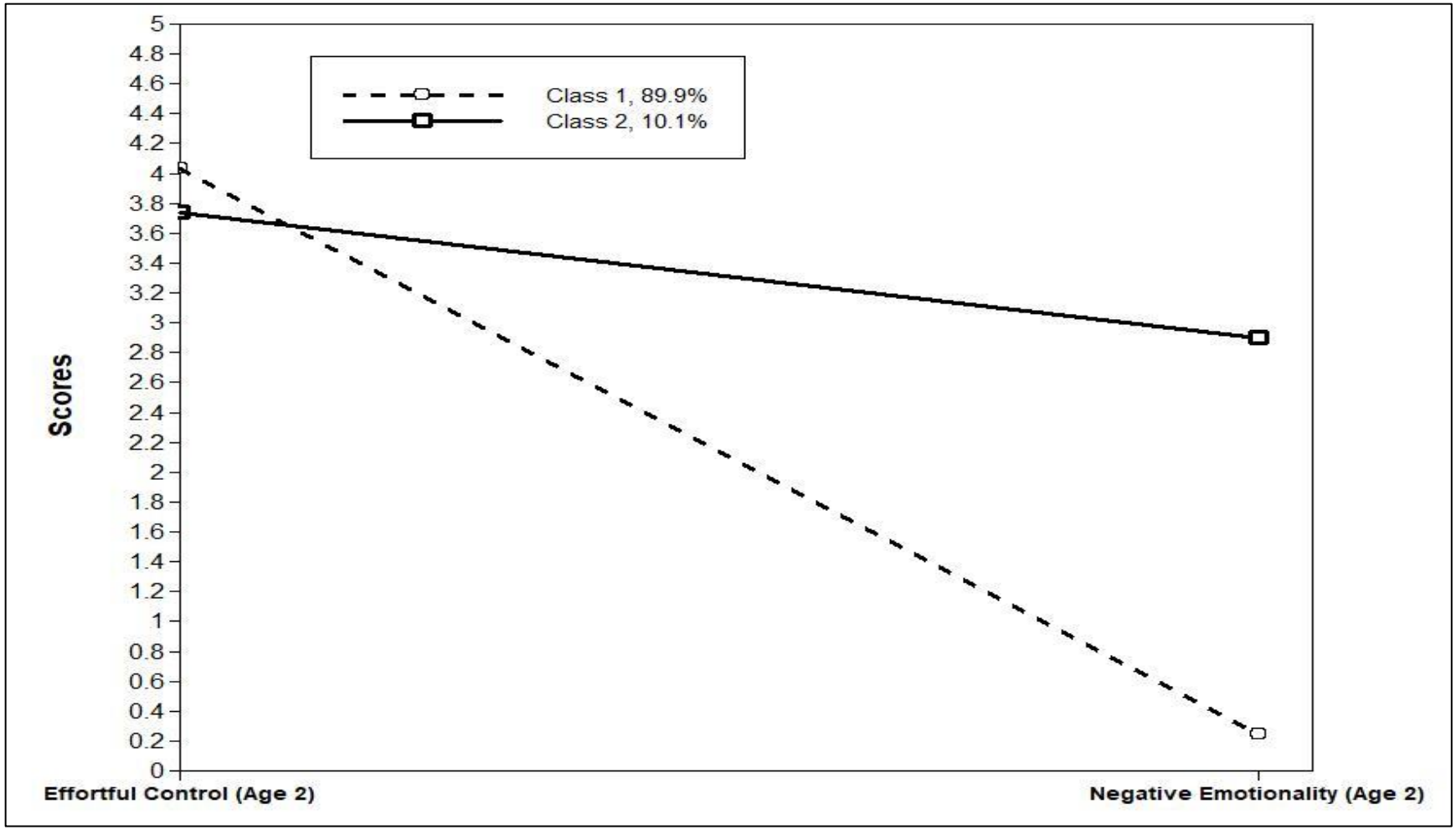

Figure 2. Estimating two latent profiles based on effortful control and negative emotionality at age 2.

3.1.3 Step 2: Classifying Children According to Their Latent Profile

The average posterior probability for each class was then calculated to confirm the adequacy of the final three-profile LPA model. Posterior probabilities measure a specific individual's likelihood of belonging to each of the model's trajectory groups (Nagin, 2005). The posterior probabilities of group membership for the Average EC/Low NE, Average EC/Moderate NE, and Average EC/High NE profiles were 99.6%, 97.5%, and 98.4% respectively. Thus, average posterior probabilities were well above recommended thresholds for assignment (i.e., > .70; Nagin, 2005), suggesting very low classification errors. Youth were then classified into the latent profile corresponding to their maximum posterior probability, and two dummy-coded variables were created to represent membership in one of the three latent groups at baseline with the Average EC/High NE profile used as the reference group. In the analyses reported next, this three-level variable was tested as a latent moderator of the association between age 3 parenting and age 10.5 CP.

3.2 AIM 3

3.2.1 Step 3: Differential Effects of Parenting by GPS and Latent Profiles of EC and NE

We next examined whether children differ in their susceptibility to positive and negative parenting in toddlerhood in relation to later CP based on their aggression GPS and their profile of phenotypic (i.e., NE and EC at age 2) characteristics identified in Aim 1. As previously noted, because of limited variability in the aggression GPS, we were unable to identify subgroups of

children with distinct genetic profiles. Thus, aggression GPS were omitted from the LPA and were instead included as a continuous moderator of parenting effects in subsequent analyses. The interaction between age 3 parenting and children's latent profile of age 2 EC/NE and the interaction between age 3 parenting and children's GPS were included in the same models predicting age 10.5 CP. Models accounted for the effects of child sex, child age in months at the age 10.5 assessments, intervention status, geographic location, family income, and ancestry principal components. Although we initially included two-way interactions between GPS and covariates, these interaction terms were removed from final models as none of them were significant and excluding them from analyses did not alter the significance of any model. Thus, we prioritized more parsimonious models including only the main effects of covariates.

Parameter estimates for models involving positive and negative parenting at age 3 are reported in Tables 4 and 5, respectively. We first present findings using the entire sample followed by race-specific findings for Caucasian and non-Caucasian (i.e., African American and Latino) youth. Although analyses were also computed for African American youth only, findings from these models were consistent with those combining African American and Latino youth together. Thus, to reduce redundancy, we prioritize the presentation of results for African American and Latino youth combined.

3.2.1.1 Positive Parenting X GPS and Latent Profiles of EC and NE (Table 4)

Whole sample analyses. For models predicting parent- and youth-reported CP at age 10.5, children's profiles of EC/NE at age 2 and positive parenting at age 3 were unrelated to age 10.5 CP. Higher aggression GPS predicted lower levels of youth-reported CP at age 10.5, but GPS were unrelated to age 10.5 CP according to parent-report. Further, neither age 2 EC/NE profiles nor

aggression GPS moderated the association between age 3 positive parenting and parent- or youth-reported CP at age 10.5.

Turning to the model predicting teacher-reported CP at age 10.5, consistent with bivariate correlations, greater positive parenting at age 3 predicted lower levels of CP in middle childhood. However, the main effect of positive parenting at age 3 on teacher-reported CP at age 10.5 was qualified by a significant interaction with age 2 EC/NE profiles. At lower levels of age 3 positive parenting, children in the Average EC/High NE profile demonstrated greater CP at age 10.5 compared to those in the Average EC/low NE profile. However, at higher levels of positive parenting at age 3, children in the Average EC/low NE and Average EC/high NE profiles showed comparable levels of teacher-reported CP. Children in the Average EC/low NE profile showed similar levels of CP at age 10.5 regardless of whether they were exposed to more or less positive parenting at age 3.

Caucasian youth. As previously noted, to address concerns of potential population stratification, interactions computed with the entire sample were followed up with race-specific analyses. Among Caucasian youth, age 2 latent profiles of EC/NE and aggression GPS were unrelated to age 10.5 CP across all informants (i.e., parents, teachers, and youth). Unexpectedly and in contrast to analyses computed with the entire sample, higher levels of positive parenting at age 3 predicted greater parent-, teacher, and youth-reported CP at age 10.5 for Caucasian youth.

Consistent with whole sample analyses, there was no evidence to suggest that aggression GPS moderated the association between age 3 positive parenting and age 10.5 CP according to any informant (i.e., parents, teachers, or youth). However, the main effect of age 3 positive parenting on age 10.5 CP was qualified by a significant interaction with age 2 EC/NE profiles when predicting parent-, teacher-, and youth-reported CP. Unexpectedly, simple slope analyses

indicated that age 3 positive parenting was positively associated with parent-reported CP for youth with Average EC/High NE, $t(30) = 2.06$, $p < .05$, but negatively associated with age 10.5 CP for youth with Average EC/Low NE, $t(30) = -2.14$, $p < .05$ (Figure 3). RoS and PoI indices (PoI = .80) suggested that individuals with Average EC/High NE and Average EC/Low NE significantly differed at average and higher levels but not lower levels of age 3 positive parenting. Positive parenting in toddlerhood was unrelated to parent-reported CP in middle childhood for youth with Average EC/Moderate NE (Figure 4), $t(30) = -1.63$, $p > .05$. The interaction between age 3 positive parenting and age 2 EC/NE profiles showed a similar pattern when predicting teacher-reported (Figures 5 and 6) and youth-reported (Figures 7 and 8) CP at age 10.5, with the exception that the slope for youth with Average EC/Low NE was non-significant in these models.

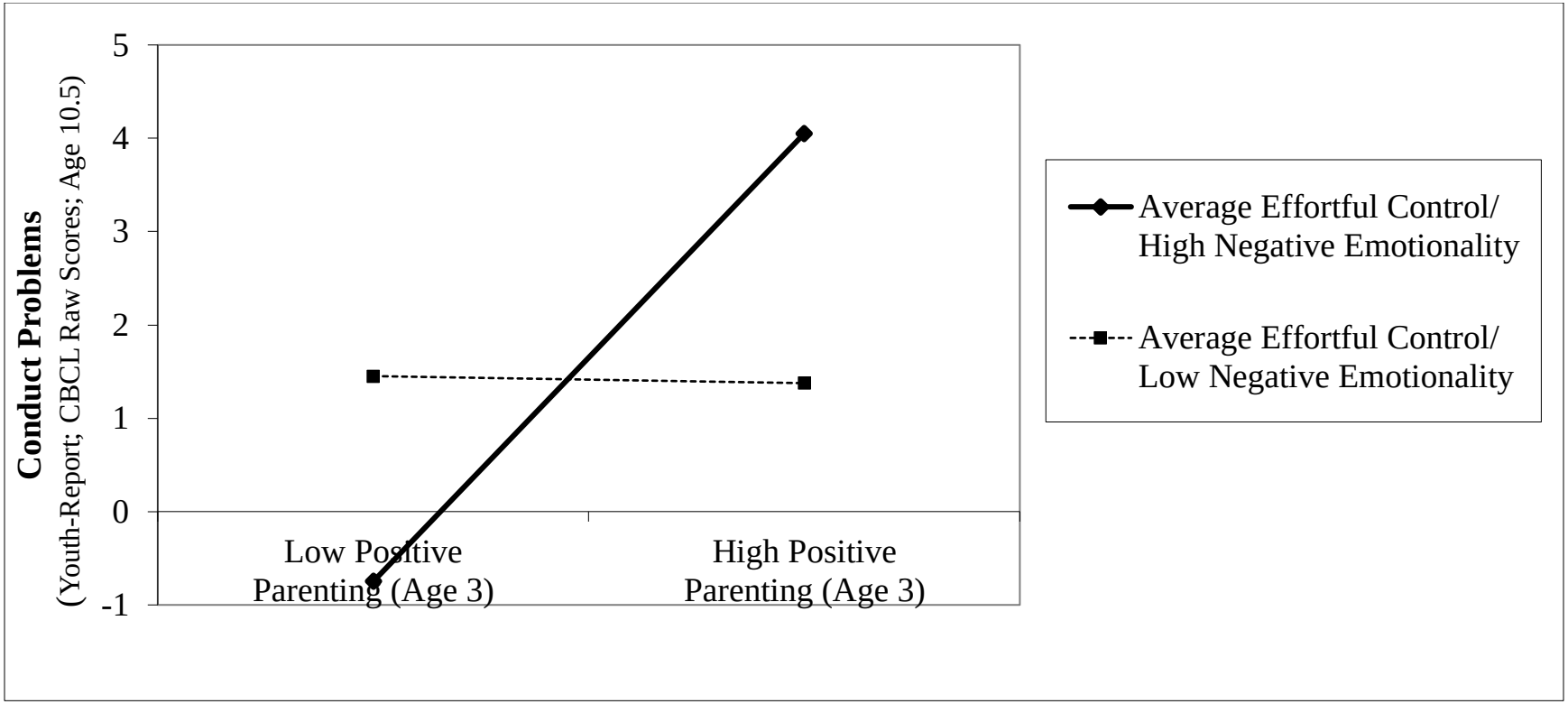

Figure 3. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = 2.06, *p* = .048. The simple slope for youth in the Average Effortful Control/Low Negative Emotionality (Avg. EC/Low NE) latent profile: *t*(30) = -2.14, *p* = 0.041. Avg. EC/High NE and Avg. EC/Low NE profiles are significantly different from each other at positive parenting values less than -1.449 and greater than 0.299. The proportion of interaction index (PoI) = 0.80.**

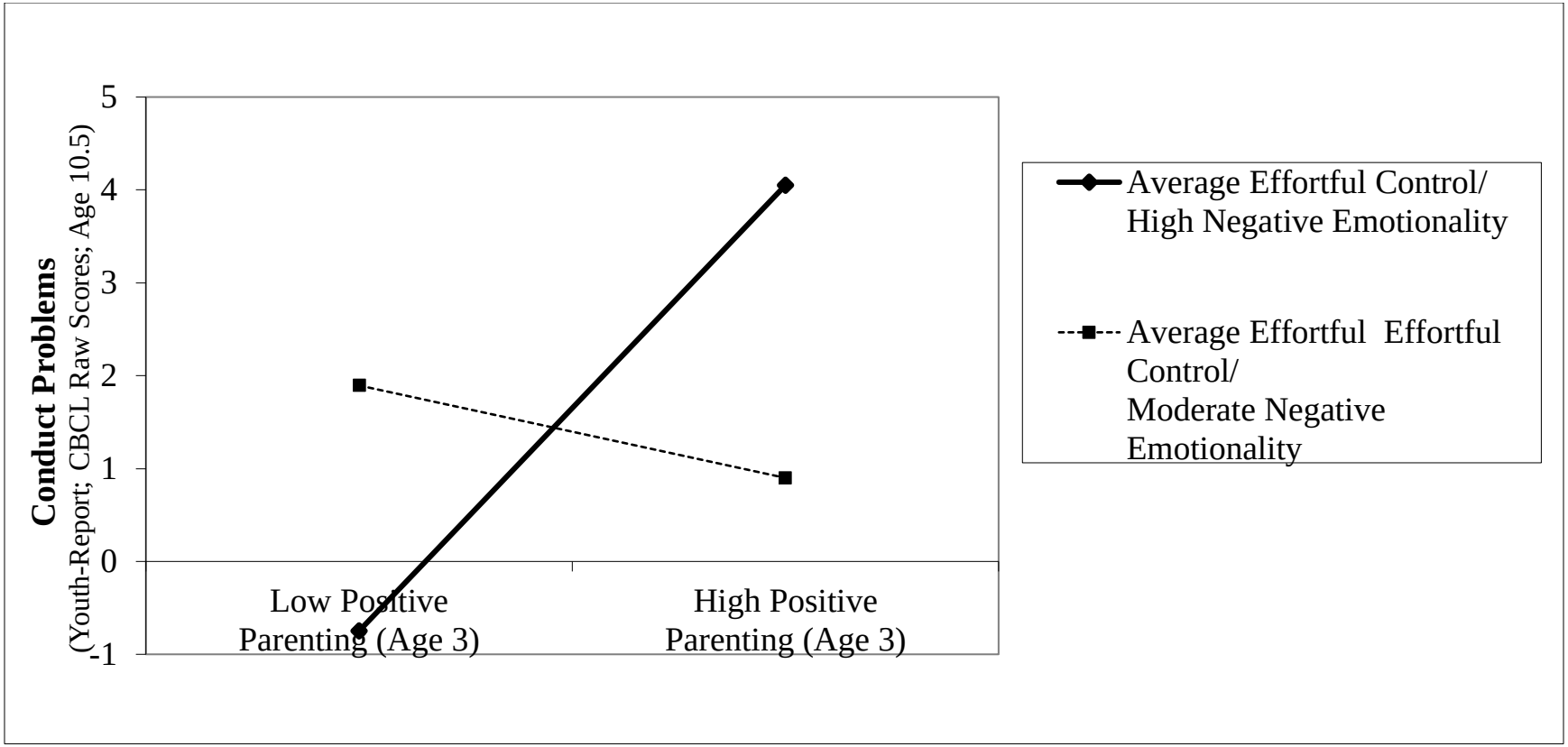

Figure 4. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = 2.06, *p* = .048. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(30) = -1.63, *p* = 0.114. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at positive parenting values less than -1.998 and greater than 0.633. The proportion of interaction index (PoI) = 0.73.**

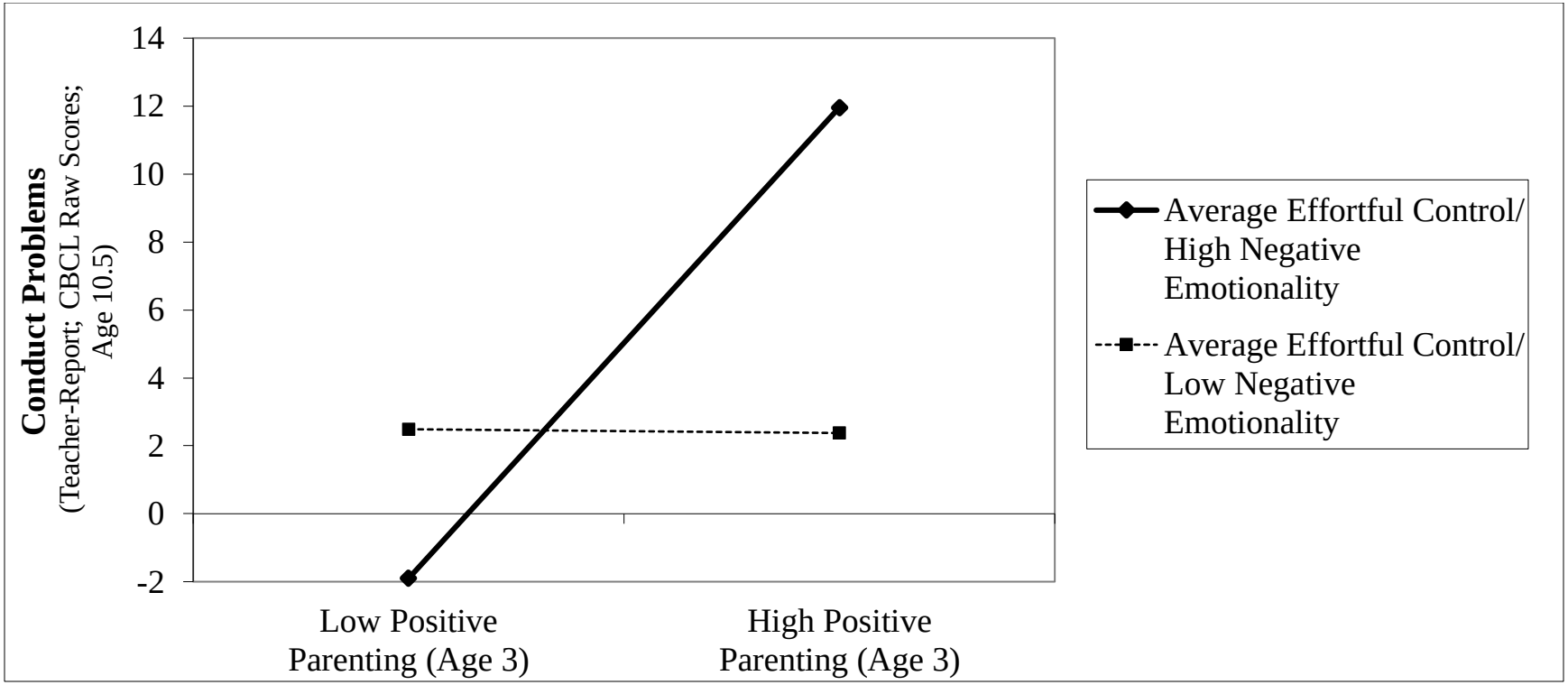

Figure 5. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (Age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = 3.78, *p* = 0.001. The simple slope for youth in the Average Effortful Control/Low Negative Emotionality (Avg. EC/Low NE) latent profile: *t*(30) = -0.15, *p* = 0.880. Avg. EC/High NE and Avg. EC/Low NE profiles are significantly different from each other at positive parenting values less than -1.008 and greater than 0.168. The proportion of interaction index (PoI) = 0.83.**

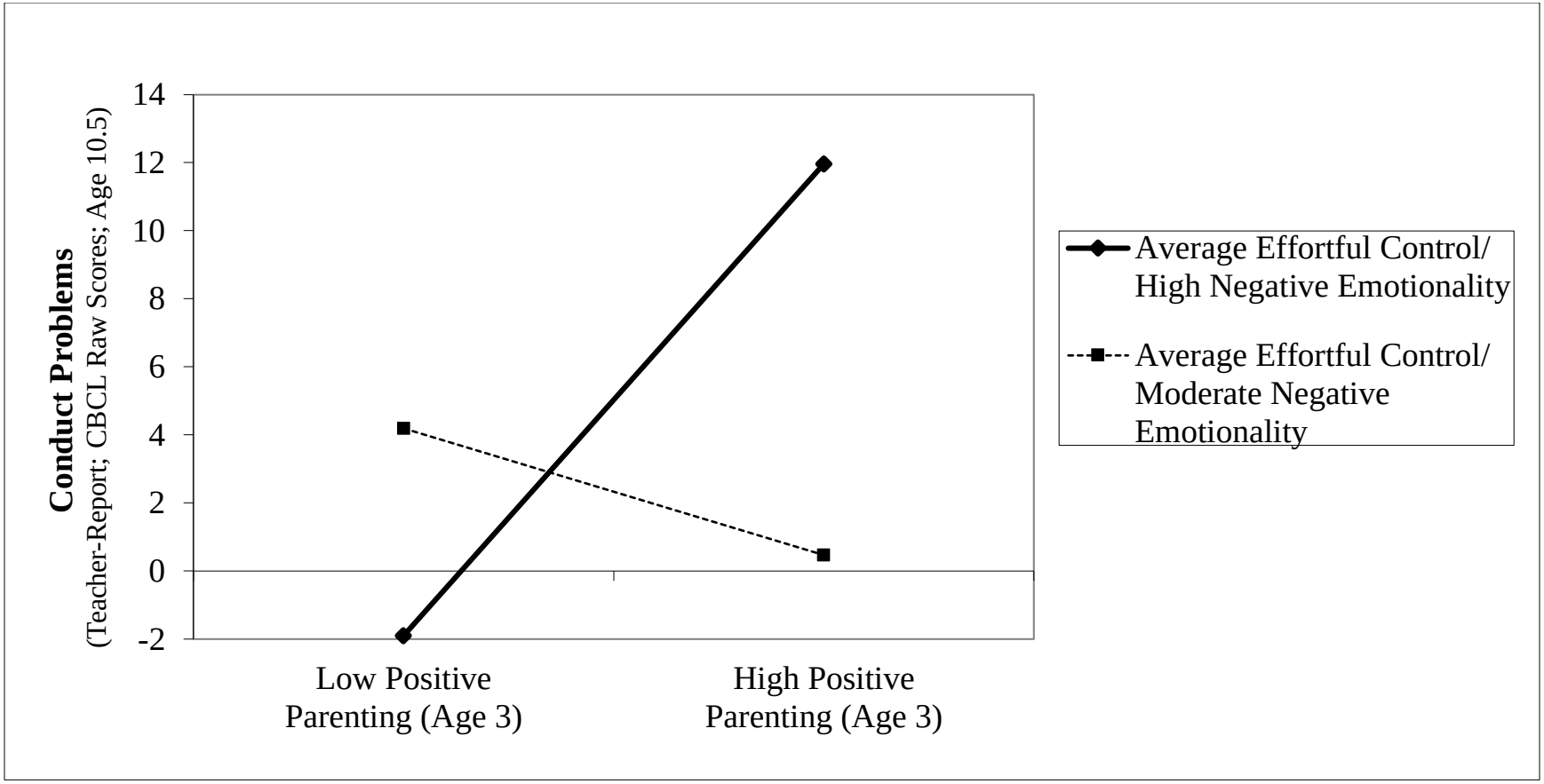

Figure 6. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = 3.78, *p* = 0.001. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(30) = -2.03, *p* = 0.051. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at positive parenting values less than -0.853 and greater than 0.144. The proportion of interaction index (PoI) = 0.78.**

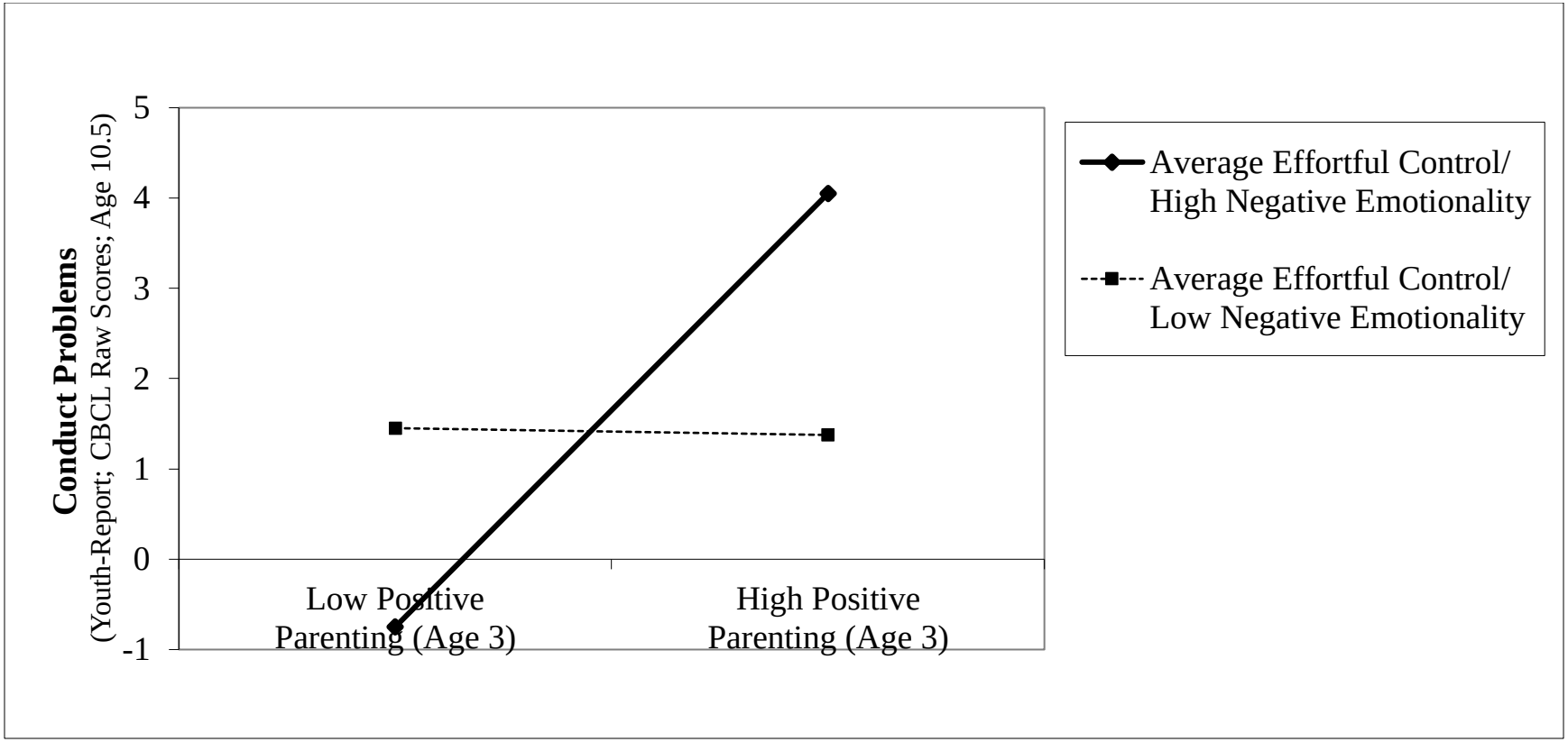

Figure 7. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = 4.53, *p* = 0.000. The simple slope for youth in the Average Effortful Control/Low Negative Emotionality (Avg. EC/Low NE) latent profile: *t*(30) = -0.22, *p* = 0.826. Avg. EC/High NE and Avg. EC/Low NE profiles are significantly different from each other at positive parenting values less than -0.549 and greater than 0.417. The proportion of interaction index (PoI) = 0.60.**

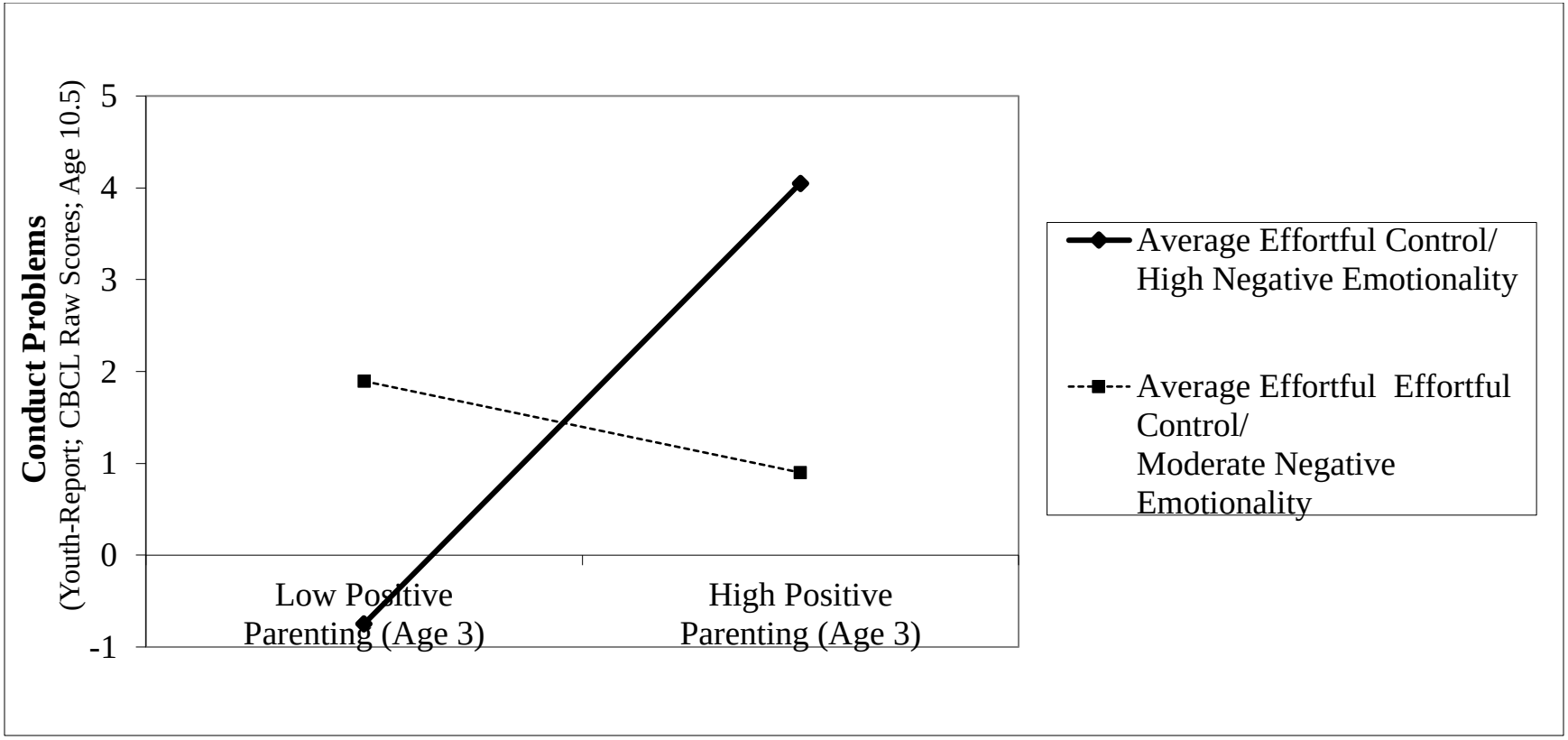

Figure 8. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = 4.53, *p* = 0.000. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(30) = -1.39, *p* = 0.176. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at positive parenting values less than -0.540 and greater than 0.365. The proportion of interaction index (PoI) = 0.59.**

Non-Caucasian youth (i.e., African American and Latino). As expected and consistent with models computed for the entire sample, for non-Caucasian youth, higher levels of age 3 positive parenting predicted lower teacher- and youth-, but not parent-reported CP at age 10.5. Further, there was a main effect of age 2 latent profiles of EC/NE on parent- and teacher-reported CP, such that youth in the Average EC/High NE profile showed greater CP compared to those in the Average EC/Moderate NE profile. While the Average EC/High profile also showed greater CP relative to the Average EC/Low NE profile, this association was only significant when CP were rated by teachers but not by parents or youth. Consistent with analyses computed with the entire sample and with Caucasian youth only, for non-Caucasian youth, aggression GPS were negatively related to youth-reported CP in middle childhood (marginal for parent-report). However, the interaction of GPS with age 3 positive parenting was non-significant for all informants of age 10.5 CP (i.e., parents, teachers, and youth). Further, children's profiles of EC/NE at age 2 moderated the association between age 3 positive parenting and teacher-reported CP at age 10.5 but not the association between positive parenting and parent- or youth-reported CP. Simple slope analyses revealed that for youth with Average EC/High NE, positive parenting at age 3 was negatively associated with teacher-reported CP at age 10.5, $t(30) = -4.64$, $p < .001$. However, positive parenting in toddlerhood was unrelated to teacher-reported CP in middle childhood for youth with Average EC/Low NE (Figure 9), $t(30) = -.91$, $p > .05$, or with Average EC/Moderate NE (Figure 9), $t(30) = -1.70$, $p > .05$. PoI indices of 0.01 and 0.02 corresponding to Figures 9 and 10, respectively, suggest that the interaction of age 3 positive parenting and age 2 EC/NE profiles was consistent with a diathesis-stress pattern. Specifically, the Average EC/High NE profile showed elevated teacher-reported CP at lower levels of age 3 positive parenting but did not significantly

differ from youth with Average EC/Low NE or Average EC/Moderate NE at higher levels of age 3 positive parenting.

Table 4. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by aggression genome-wide polygenic scores (GPS) and age 2 latent profiles of effortful control (EC) and negative emotionality (NE)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.135 (0.047)[1]	0.004	0.186 (0.064)[1,2]	0.004	0.006 (0.050)	0.906
Child Age	-0.062 (0.046)[1]	0.173	-0.002 (0.048)[1]	0.964	0.073 (0.048)	0.127
Intervention Status	0.085 (0.049)	0.082	0.012 (0.046)	0.789	0.037 (0.050)	0.461
Pittsburgh Site	0.047 (0.065)	0.476	0.078 (0.058)[2]	0.184	0.158 (0.058)[2]	0.006
Virginia Site	-0.097 (0.050)[1]	0.052	0.021 (0.055)	0.698	0.030 (0.059)	0.610
Family Income	-0.040 (0.059)	0.497	-0.072 (0.053)[1]	0.173	-0.086 (0.036)[2]	0.018
Ancestry PC1	0.026 (0.066)	0.691	0.138 (0.070)	0.047	0.115 (0.060)	0.057
Ancestry PC2	0.030 (0.041)	0.465	0.023 (0.037)	0.537	0.024 (0.036)	0.495
Average EC/Low NE (Age 2)[a]	0.020 (0.065)	0.758	-0.210 (0.106)[2]	0.047	-0.062 (0.073)	0.396
Average EC/Moderate NE (Age 2)[a]	0.116 (0.080)[2]	0.148	-0.153 (0.095)[2]	0.107	0.068 (0.083)	0.413
GPS	-0.063 (0.053)[2]	0.232	-0.106 (0.053)	0.047	-0.139 (0.059)[2]	0.019
Pos. Parenting (Age 3)	0.128 (0.145)[1]	0.378	-0.492 (0.194)[1,2]	0.011	-0.181 (0.161)[1,2]	0.261
Pos. Parenting X Avg. EC/Low NE	-0.222 (0.138)[1]	0.109	0.431 (0.187)[1,2]	0.021	0.152 (0.163)[1]	0.352
Pos. Parenting X Avg. EC/Moderate NE	-0.027 (0.089)[1]	0.759	0.078 (0.098)[1,2]	0.429	-0.005 (0.083)[1]	0.952
GPS X Pos. Parenting	0.022 (0.054)	0.691	0.024 (0.046)	0.601	0.083 (0.059)	0.159

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity; GPS = genome-wide polygenic scores. Model parameters reported separately by race can be found in the Appendix available at the end of the document.**

[a] The Average EC/High NE latent profile is the reference group.

[1] Significant for Caucasian participants. [1] Marginally significant for Caucasian participants. [2] Significant for non-Caucasian participants. [2] Marginally significant for non-Caucasian participants.

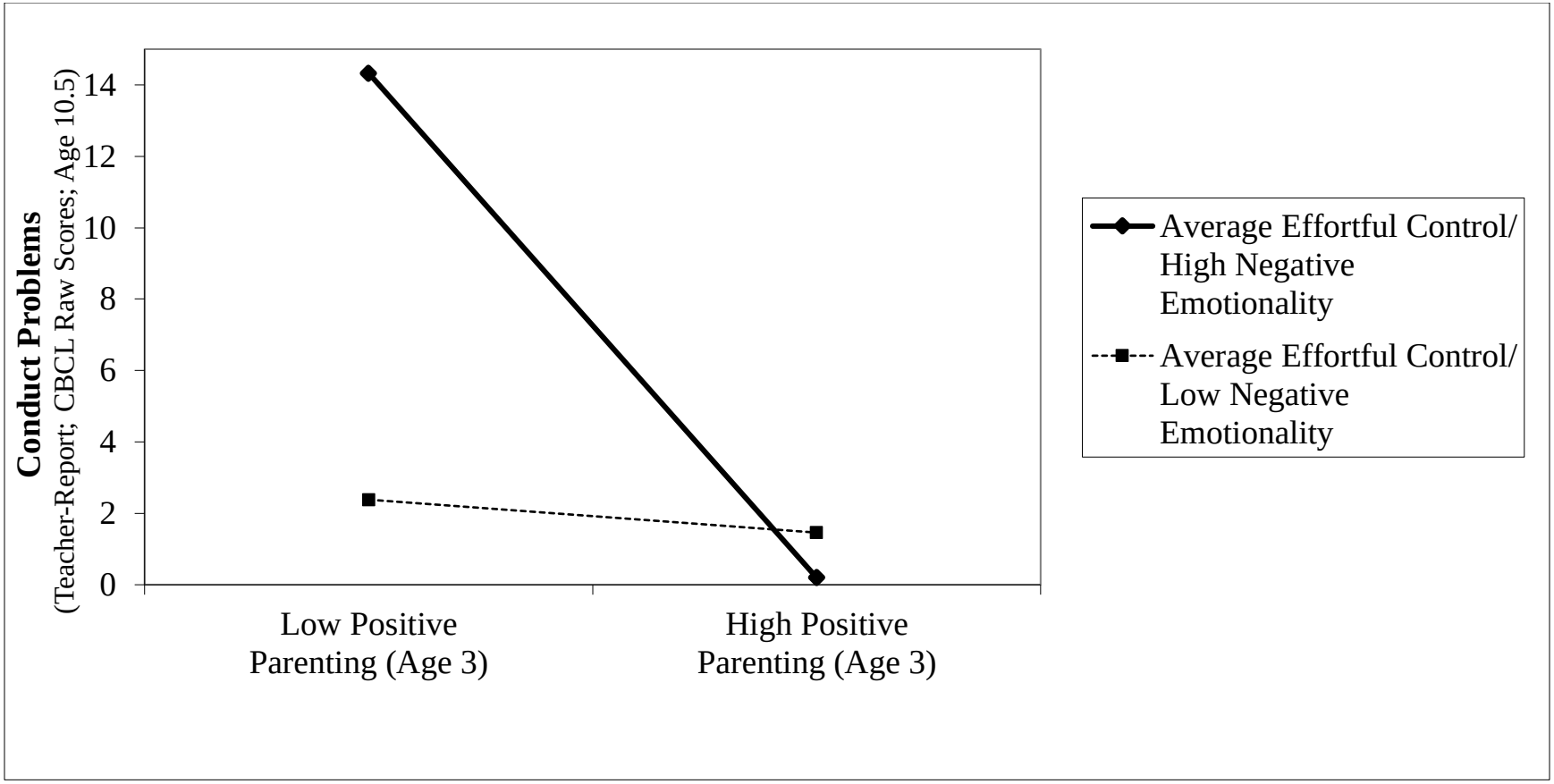

Figure 9. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Non-Caucasian youth only.

***Note*. The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: $t(30) = -4.64$, $p = 0.000$. The simple slope for youth in the Average Effortful Control/Low Negative Emotionality (Avg. EC/Low NE) latent profile: $t(30) = -0.91$, $p = 0.372$. Avg. EC/High NE and Avg. EC/Low NE profiles are significantly different from each other at positive parenting values less than 0.342 and greater than 1.535. The proportion of interaction index (PoI) = 0.01.**

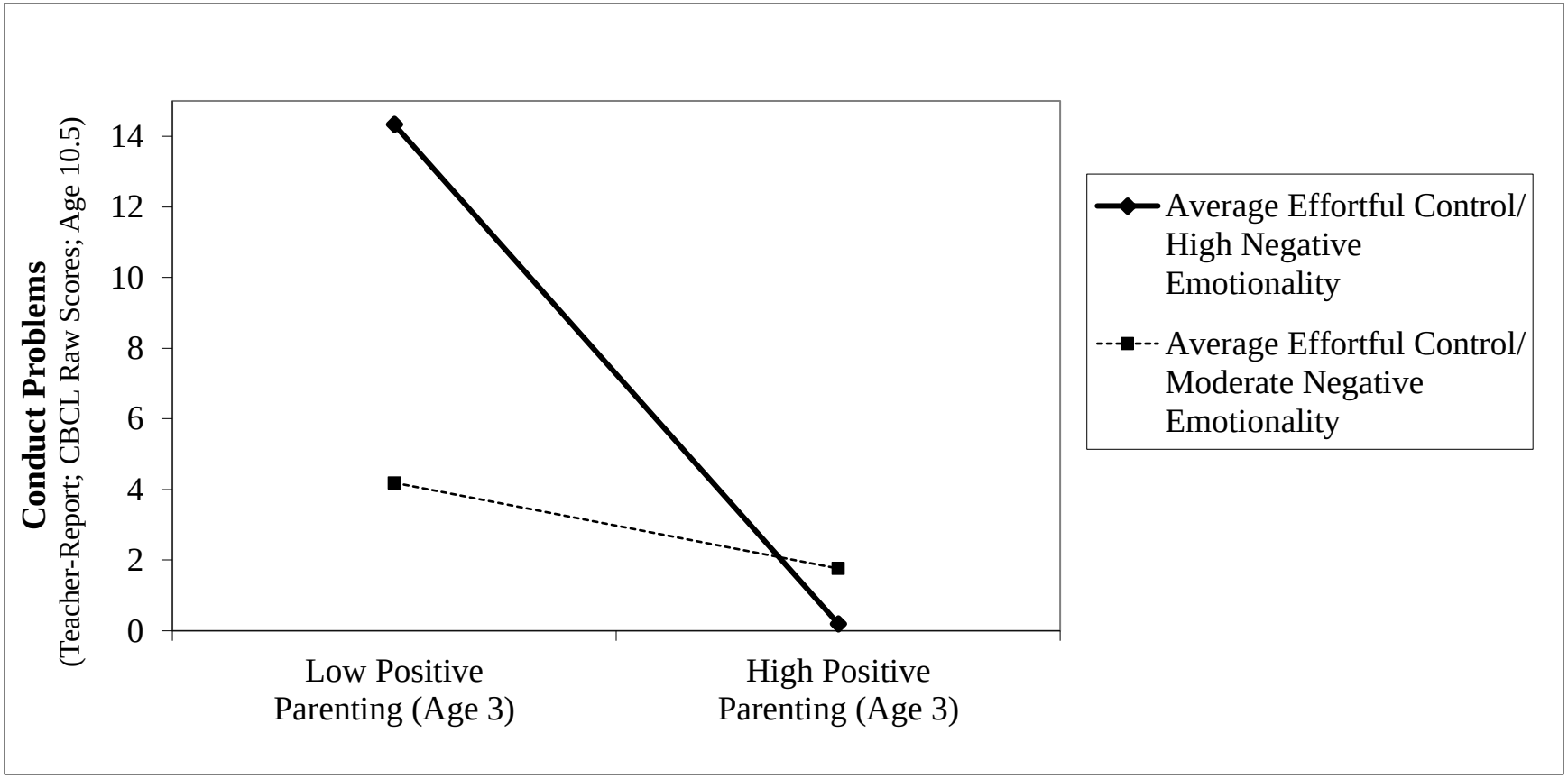

Figure 10. Positive parenting (age 3) x latent profiles of effortful control and negative emotionality (Age 2) for Non-Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile *t*(30) = -4.64, *p* = 0.000. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(30) = -1.70, *p* = 0.100. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at positive parenting values less than -0.011 and greater than 1.726. The proportion of interaction index (PoI) = 0.02.**

3.2.1.2 Negative Parenting X GPS and Latent Profiles of EC and NE (Table 5)

Whole sample analyses. The same models were recomputed replacing positive parenting at age 3 with negative parenting at age 3. Children's profiles of EC/NE at age 2, negative parenting at age 3, and aggression GPS were unrelated to parent-reported CP at age 10.5. Further, neither age 2 EC/NE profiles nor aggression GPS moderated the association between age 3 negative parenting and parent-reported CP at age 10.5.

Turning to the model predicting teacher-reported CP at age 10.5, age 2 EC/NE profiles and age 3 negative parenting were unrelated to CP in middle childhood. While higher aggression GPS predicted lower teacher-reported CP, consistent with models predicting parent- and youth-reported CP, there was no evidence to suggest that aggression GPS moderated the association between positive parenting at age 3 and teacher-reported CP at age 10.5 However, children's EC/NE profiles at age 2 moderated the association between negative parenting at age 3 and teacher-reported CP at age 10.5. As shown in Figure 11, for children in the Average EC/High NE profile, less negative parenting at age 3 was related to more CP at age 10.5. In contrast, for children in the Average EC/Moderate NE profile, less negative parenting at age 3 predicted fewer teacher-reported CP at age 10.5. Children in the Average EC/High NE profile and the Average EC/Moderate NE profile showed similar levels of age 10.5 CP when exposed to high levels of negative parenting at age 3. The interaction between age 3 negative parenting and age 2 EC/NE on age 10.5 CP was consistent with a vantage-sensitivity pattern as indicated by RoS and PoI indices (PoI = .05).

Caucasian youth. We then tested the interaction between age 3 negative parenting and aggression GPS and age 2 EC/NE profiles among Caucasian youth only. Consistent with models computed for the entire sample, age 2 EC/NE profiles and age 3 negative parenting were unrelated

to age 10.5 CP across all informants (i.e., parents, teachers, and youth) for Caucasian youth. While aggression GPS exerted direct and interactive effects on age 10.5 CP in the entire sample, these effects were absent when analyses were computed with Caucasian youth. However, age 2 EC/NE profiles moderated the association between age 3 negative parenting and youth-reported CP for Caucasian youth. Simple slope analyses indicated that for youth with Average EC/Moderate NE, negative parenting at age 3 was positively associated with youth CP at age 10.5 (Figure 12), $t(30) = 2.48$, $p < .05$. However, negative parenting in toddlerhood was unrelated to youth-reported CP in middle childhood for youth with Average EC/High NE, $t(30) = -1.33$, $p > .05$. Specifically, youth with Average EC/Moderate NE showed reduced CP when exposed to lower levels of negative parenting at age 3 but did not significantly differ from youth with Average EC/High NE when exposed to average or high levels of negative parenting.

Table 5. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by aggression genome-wide polygenic scores (GPS) and age 2 latent profiles of effortful control and negative emotionality

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.141 (0.047)[1,2]	0.002	0.211 (0.077)[2]	0.006	0.011 (0.052)	0.838
Child Age	-0.077 (0.042)[1]	0.071	-0.029 (0.051)	0.573	0.043 (0.048)	0.363
Intervention Status	0.077 (0.049)[1]	0.113	0.007 (0.051)	0.893	0.038 (0.049)	0.441
Pittsburgh Site	0.061 (0.064)	0.346	0.08 (0.062)[2]	0.198	0.153 (0.057)[1,2]	0.008
Virginia Site	-0.094 (0.05)[1]	0.058	0.013 (0.063)	0.840	0.017 (0.059)	0.777
Family Income	-0.036 (0.058)	0.538	-0.084 (0.059)	0.154	-0.085 (0.036)[2]	0.020
Ancestry PC1	0.024 (0.064)	0.711	0.148 (0.079)	0.060	0.109 (0.060)	0.068
Ancestry PC2	0.031 (0.041)	0.448	0.029 (0.044)	0.510	0.022 (0.036)	0.537
Avg. EC/Low NE (Age 2)[a]	0.009 (0.064)	0.887	-0.162 (0.141)	0.251	-0.054 (0.068)	0.434
Avg. EC/Moderate NE (Age 2)[a]	0.096 (0.080)[2]	0.230	-0.138 (0.122)	0.256	0.058 (0.070)	0.409
GPS	-0.061 (0.052)	0.240	-0.106 (0.057)	0.064	-0.126 (0.049)[2]	0.009
Neg. Parenting (Age 3)	-0.16 (0.172)	0.350	-0.247 (0.354)	0.486	0.113 (0.229)[1]	0.623
Neg. Parenting X Avg. EC/Low NE [a]	0.185 (0.160)	0.247	0.333 (0.301)	0.267	-0.083 (0.223)[1]	0.712
Neg. Parenting X Avg. EC/Moderate NE [a]	0.122 (0.076)	0.110	0.235 (0.090)[2]	0.009	0.109 (0.102)[1]	0.287
GPS X Neg. Parenting	-0.075 (0.049)[2]	0.128	-0.044 (0.046)	0.335	-0.129 (0.066)[2]	0.054

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC1 and 2 are principle components accounting for genetic variation due to race/ethnicity, GPS = genome-wide polygenic scores. Model parameters reported separately by race can be found in the Appendix available at the end of the document.**

[a] The Average EC/High NE latent profile is the reference group.

[1] Significant for Caucasian participants. [1] Marginally significant for Caucasian participants. [2] Significant for non-Caucasian participants. [2] Marginally significant for non-Caucasian participants.

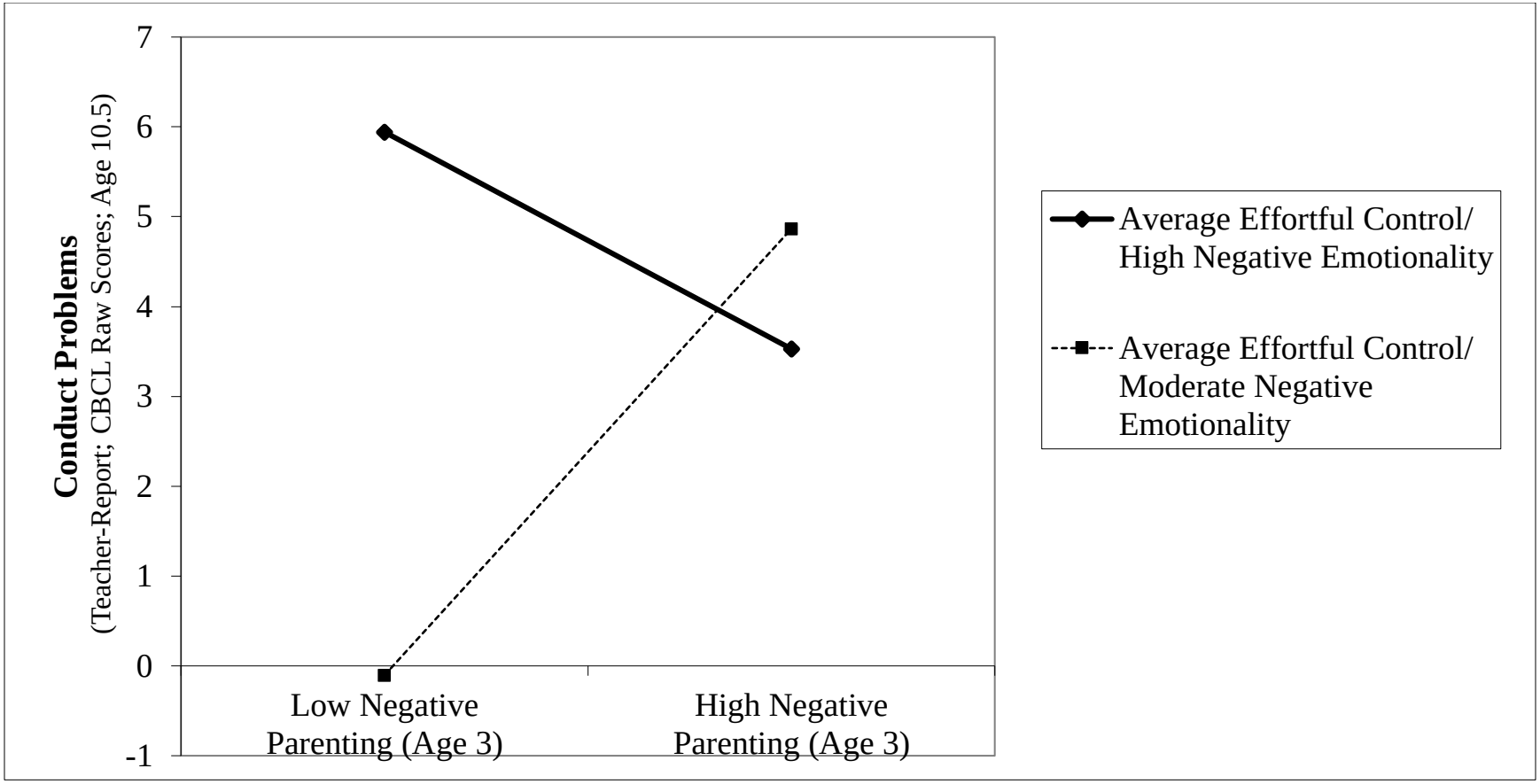

Figure 11. Negative parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for whole sample.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(36) = -0.59, *p* = 0.560. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(36) = 2.96, *p* = 0.005. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at negative parenting values less than -2.223 and greater than -0.304. The proportion of interaction index (PoI) = 0.05.**

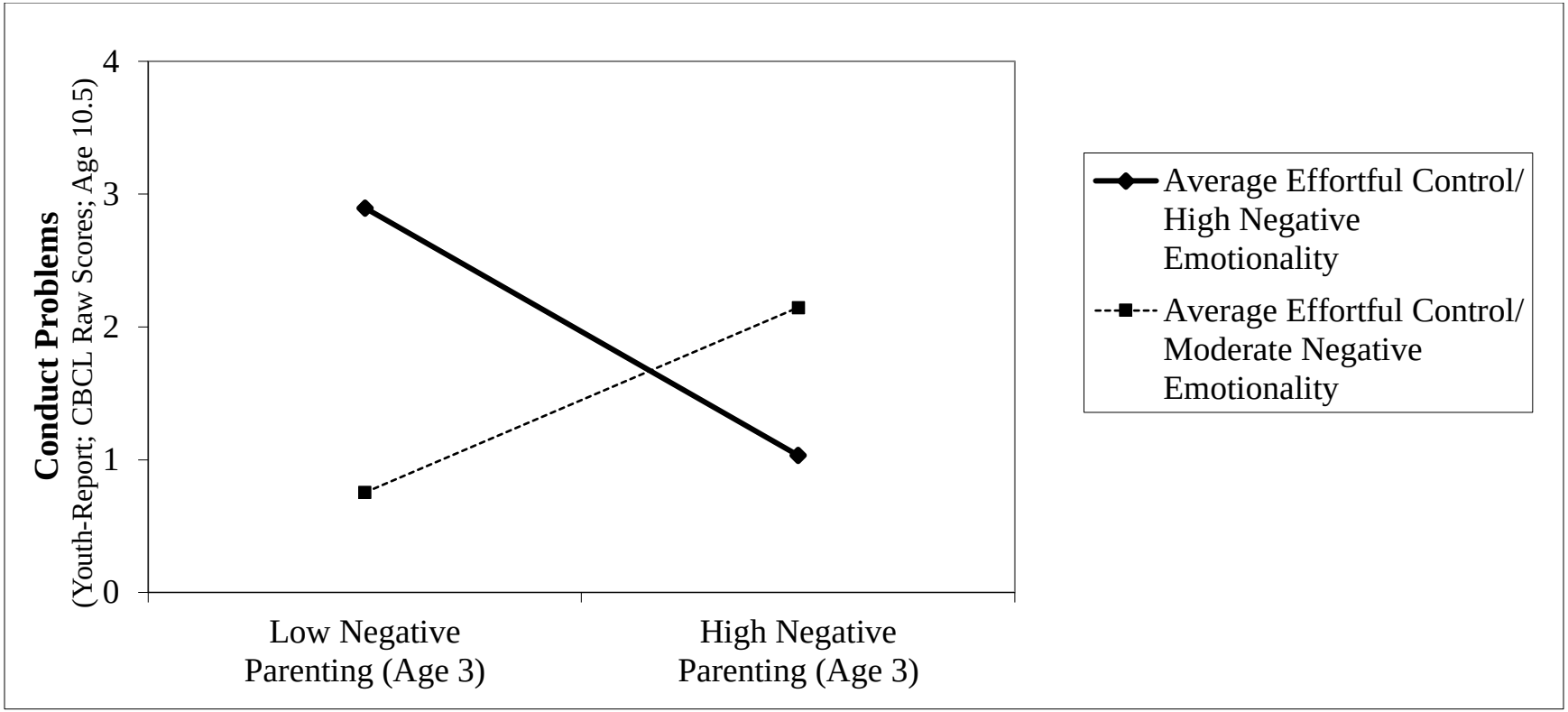

Figure 12. Negative parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = -1.33, *p* = 0.192. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(30) = 2.48, *p* = 0.019. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at negative parenting values less than -0.375 and greater than 4.528. The proportion of interaction index (PoI) = 0.21.**

Non-Caucasian youth (i.e., African American and Latino). As expected and consistent with analyses computed with the entire sample, age 3 observed negative parenting and age 2 EC/NE profiles were unrelated to age 10.5 CP across all informants for non-Caucasian youth. Also consistent with whole sample analyses, aggression GPS were significantly associated with youth, but not parent- or teacher-reported CP at age 10.5. Further, the interaction between age 3 negative parenting and age 2 EC/NE profiles in relation to teacher-reported CP at age 10.5 that we identified in the entire sample was replicated with non-Caucasian youth. For youth with Average EC/Moderate NE at age 2, negative parenting at age 3 was positively associated with youth CP at age 10.5 (Figure 13), $t(30) = 2.87$, $p < .01$. However, negative parenting in toddlerhood was unrelated to teacher-reported CP in middle childhood for non-Caucasian youth with Average EC/High NE, $t(30) = -0.90$, $p > .05$. RoS and PoI indices (PoI = .43) suggested that the interaction of age 3 negative parenting with age 2 EC/NE in relation to age 10.5 CP was consistent with differential susceptibility. Specifically, youth with Average EC/Moderate NE showed reduced age 10.5 CP compared to youth with Average EC/High NE when exposed to lower levels of negative parenting at age 3 but significantly greater CP when exposed to higher levels of negative parenting.

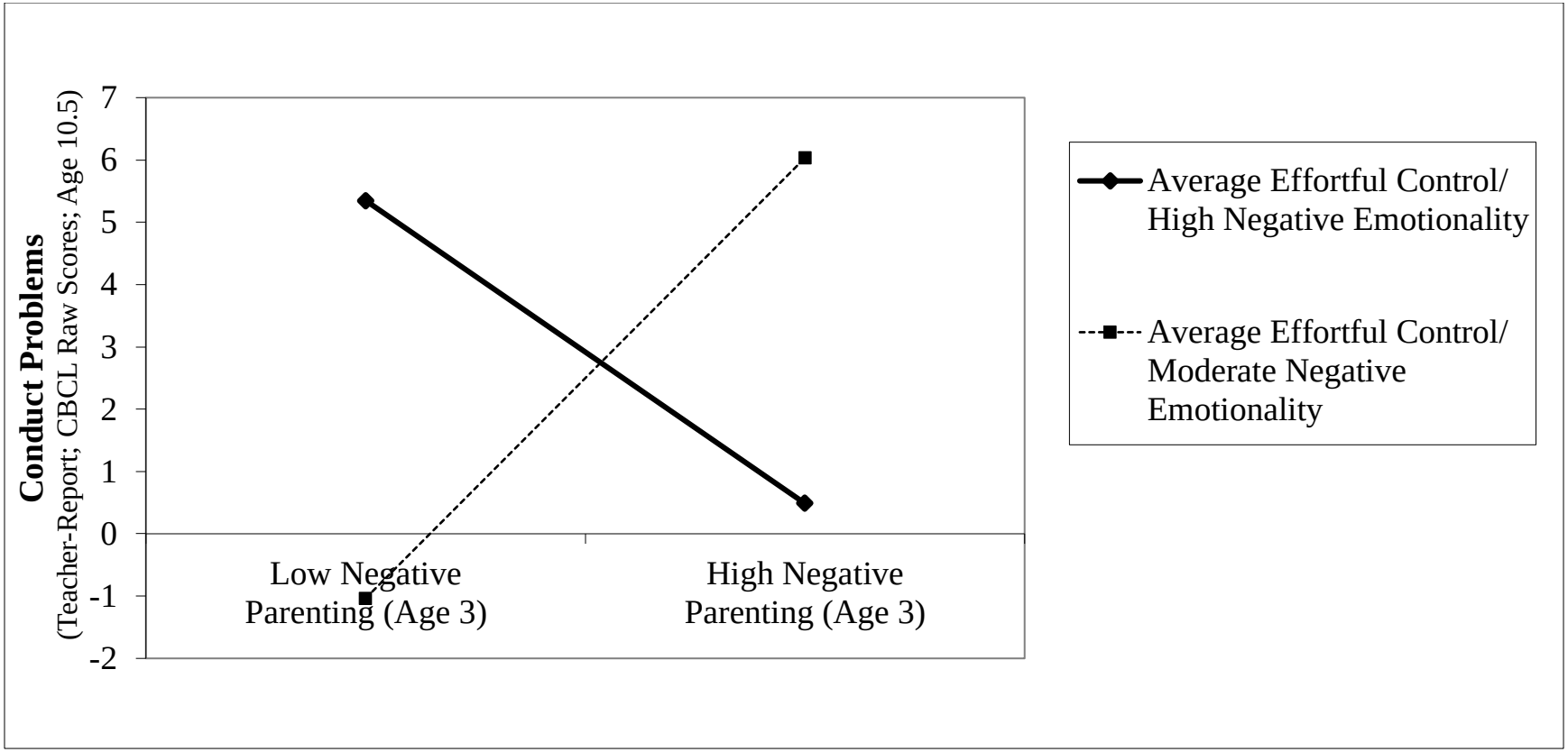

Figure 13. Negative parenting (age 3) x latent profiles of effortful control and negative emotionality (age 2) for Non-Caucasian youth only.

Note. **The simple slope for youth in the Average Effortful Control/High Negative Emotionality (Avg. EC/High NE) latent profile: *t*(30) = -0.90, *p* = 0.374. The simple slope for youth in the Average Effortful Control/Moderate Negative Emotionality (Avg. EC/Moderate NE) latent profile: *t*(30) = 2.87, *p* = 0.007. Avg. EC/High NE and Avg. EC/Moderate NE profiles are significantly different from each other at negative parenting values less than -18.657 and greater than -0.694. The proportion of interaction index (PoI) = 0.43.**

3.3 AIM 2

3.3.1 Descriptive Statistics and Zero-Order Correlations

Bivariate correlations were computed to test the hypothesis that there are moderate associations between child negative emotional (NE) reactivity at age 2 and positive emotional (PE) reactivity at age 2 (see Table 2). Consistent with Hypothesis 2, children's observed NE reactivity at age 2 was strongly correlated with observed PE reactivity scores at age 2 (r = .587). Unexpectedly, neither NE nor PE reactivity at age 2 were associated with age 10.5 CP according to parent-, teacher-, or youth-report.

3.3.2 Step 1: Identifying Latent Profiles of Emotional Reactivity

Next, to test the hypothesis that a minority of children are differentially reactive to their parent's display of positive and negative emotions at age 2, LPA was implemented to identify subsets of children who show similar profiles of emotional reactivity (e.g., high PE reactivity and high NE reactivity; low PE reactivity and high NE reactivity). As shown in Table 6, AIC, BIC, adjusted BIC, and BLRT values supported a four-profile solution. The significant bootstrap likelihood ratio test (BLRT) also provided converging evidence suggesting improved model fit when estimating four versus three classes. While the five-profile model demonstrated lower AIC, BIC, and adjusted BIC scores compared to the four-profile model, the five-profile model resulted

in a group that only consisted of 2.4% participants. Thus, the four-profile solution was retained as the best fitting model.

The resulting four profiles of emotional reactivity at age 2 are depicted in Figure 14. Profile 1 consisted of 5.4% of participants. Children in this profile displayed moderate levels of NE reactivity and high levels of PE reactivity. This subgroup is henceforth referred to as the Moderate NE/High PE profile. Profile 2 consisted of 13.6% of participants and showed moderate levels of NE reactivity and low levels of PE reactivity (i.e., Moderate NE/Low PE profile). The largest subgroup, Profile 3 (75.9%), was characterized by low levels of NE and PE reactivity (i.e., Low NE/Low PE profile), while the smallest subgroup, Profile 4 (5.2%) showed very high levels of NE and PE emotional reactivity (i.e., High NE/High PE profile).

Table 6. Aim 2: Fit indices for one to five group latent profile models based on emotional reactivity scores at age 2

# Profiles	Loglikelihood	AIC	BIC	Adjusted BIC	Entropy	*p-value* BLRT	Class Sizes (%)
1	-1626.126	3260.251	3277.118	3264.421	N/A[a]	N/A[a]	100
2	-1433.769	2881.538	2911.054	2888.835	0.923	*p* <.001	13.83/86.17
3	-1357.601	2735.201	2777.367	2745.627	0.942	*p* <.001	18.11/5.73/76.16
4	**-1313.933**	**2653.867**	**2708.682**	**2667.420**	**0.951**	***p* <.001**	**5.38/13.58/75.87/5.17**
5	-1247.961	2527.921	2595.387	2544.602	0.990	*p* <.001	13.74/2.40/21.55/56.92/5.39

***Note*. AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; BLRT = Bootstrap Likelihood Ratio Test for k versus k-1 classes; Bolded text indicates best fitting model chosen.**

[a] Entropy and BLRT are not available for one-class models.

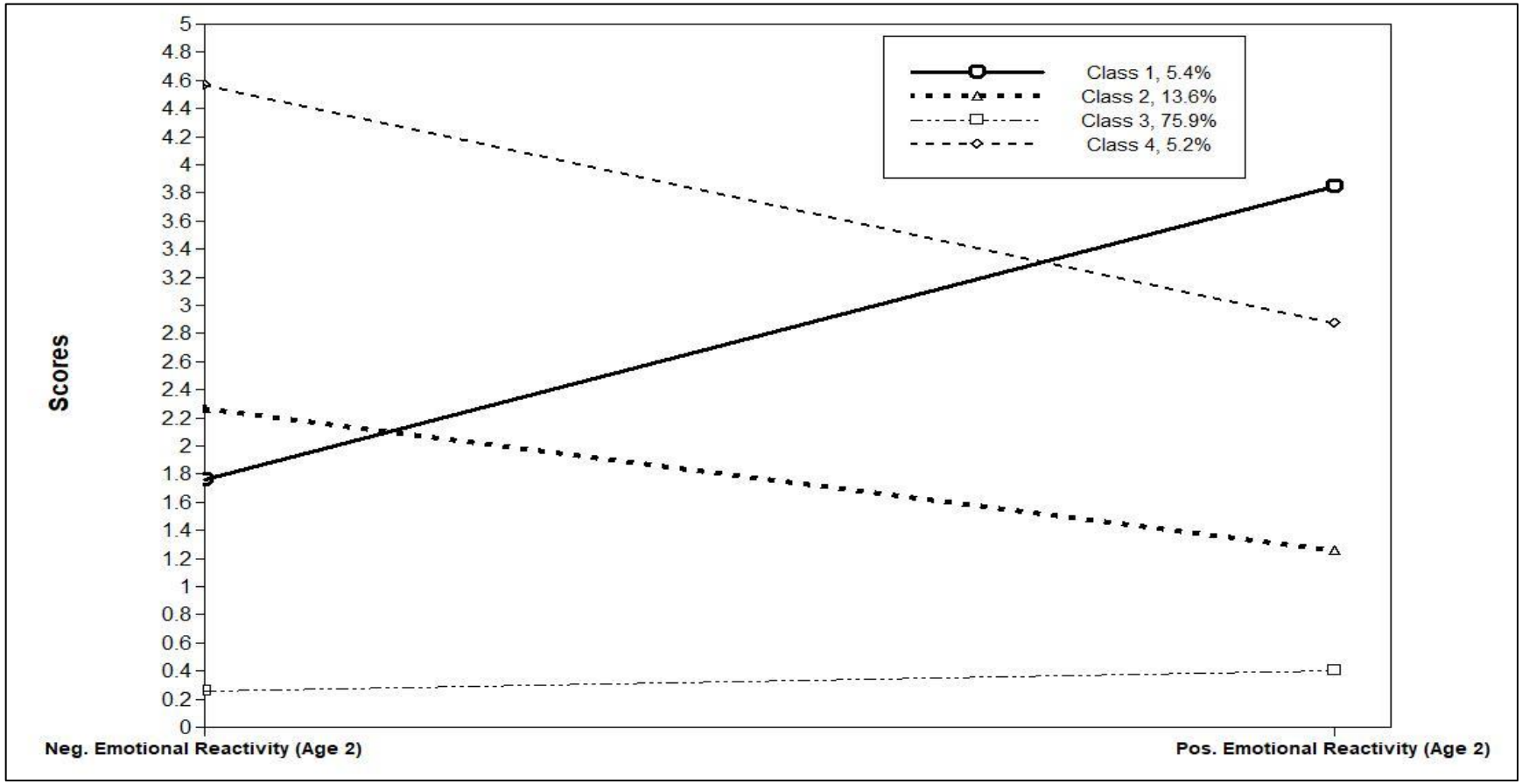

Figure 14. Latent profiles of positive and negative emotional reactivity at age 2.

Note. **Class 1 = Moderate Negative Emotional Reactivity/High Positive Emotional Reactivity; Class 2 = Moderate Negative Emotional Reactivity/Low Positive Emotional Reactivity; Class 3 = Low Negative Emotional Reactivity/Low Positive Emotional Reactivity; Class 4 = High Negative Emotional Reactivity/High Positive Emotional Reactivity.**

3.3.3 Step 2: Classifying Children According to Their Latent Profile

The average posterior probability for each class was then calculated to confirm the adequacy of the final four-profile model. The posterior probabilities of group membership for the Moderate NE/High PE profile, Moderate NE/Low PE profile, Low NE/Low PE profile, and High NE/High PE profile groups were 90.6%, 96.5%, 98.9%, and 97.6%, respectively. Thus, for all four profiles, the average posterior probabilities were all above recommended thresholds (Nagin, 2005), suggesting reasonably low classification errors. Youth were then classified into the latent profile corresponding to their maximum posterior probability. Three dummy-coded variables were created to represent membership in one of the four latent profiles at baseline (with the Low NE/Low PE profile as the reference group). This four-level variable was then used to represent the latent moderator in the moderation analyses reported next.

3.4 AIM 4

3.4.1 Step 3: Differential Effects of Parenting Across Latent Profiles of Emotional Reactivity

To examine the extent to which children's NE and PE reactivity assessed at a micro, moment-to-moment scale translates into differential susceptibility across longer periods of development, we tested the interaction between children's emotional reactivity profiles at age 2 and their exposure to positive and negative parenting at age 3 in relation to age 10.5 CP. In addition

to estimating the main effects of age 2 emotional reactivity profiles, the mains effect of age 3 parenting, and the interaction between age 2 emotional reactivity and age 3 parenting, models also accounted for the effects of child sex, child age in months at the age 10.5 assessments, intervention status, geographic location, family income, and ancestry principal components. Parameter estimates for models involving positive and negative parenting at age 3 are reported in Tables 7 and 8, respectively.

3.4.1.1 Positive Parenting X Latent Profiles of Emotional Reactivity (Table 7)

Consistent with bivariate correlations, positive parenting at age 3 negatively predicted parent-reported CP at age 10.5. Positive parenting at age 3 was unrelated to teacher- or youth-reported CP. For models predicting parent- and youth-reported CP, children's emotional reactivity profiles at age 2 were unrelated to age 10.5 CP and there was no evidence of an interaction between emotional reactivity at age 2 and positive parenting at age 3 in relation to youth CP. However, emotional reactivity at age 2 was related to teacher-report of CP, such that children in the High NE/High PE profile showed higher levels of CP at age 10.5 compared to youth in the Low NE/Low PE profile. While there was no main effect of positive parenting at age 2 on age 10.5 CP according to teacher-report, consistent with Hypothesis 4, positive parenting interacted with emotional reactivity to predict teacher-reported CP. As shown in Figure 15, compared to children in the Low NE/Low PE profile, children in the High NE/High PE profile showed significantly higher levels of teacher-reported CP at age 10.5 when exposed to low levels of positive parenting at age 3. However, at high levels of positive parenting at age 3, the Low NE/Low PE and High NE/High PE profiles showed comparable levels of age 10.5 CP according to teacher report.

Table 7. Aim 4: Age 2 emotional reactivity profiles by age 3 positive parenting predicting age 10.5 conduct problems (CP)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.151 (0.045)	0.001	0.255 (0.054)	0.000	0.066 (0.052)	0.208
Intervention Status	0.073 (0.048)	0.128	-0.038 (0.057)	0.505	0.015 (0.051)	0.764
Pittsburgh Site	0.059 (0.066)	0.371	0.113 (0.071)	0.114	0.096 (0.056)	0.088
Virginia Site	-0.062 (0.052)	0.234	0.063 (0.074)	0.396	-0.002 (0.059)	0.968
Family Income	-0.042 (0.059)	0.474	-0.115 (0.061)	0.057	-0.096 (0.037)	0.009
Ancestry PC1	0.014 (0.061)	0.815	0.125 (0.075)	0.099	0.092 (0.055)	0.093
Ancestry PC2	0.044 (0.036)	0.218	0.057 (0.041)	0.165	-0.001 (0.035)	0.983
Moderate NE/High PE [a]	-0.035 (0.040)	0.371	0.019 (0.044)	0.667	-0.020 (0.047)	0.665
Moderate NE/Low PE [a]	-0.001 (0.054)	0.985	-0.013 (0.055)	0.810	-0.011 (0.039)	0.784
High NE/High PE [a]	0.032 (0.053)	0.539	0.160 (0.070)	0.022	0.073 (0.044)	0.102
Pos. Parenting (Age 3)	-0.107 (0.057)	0.061	0.010 (0.077)	0.926	-0.016 (0.057)	0.779
Pos. Parenting X Moderate NE/High PE [a]	-0.001 (0.033)	0.974	-0.046 (0.038)	0.225	0.019 (0.020)	0.336
Pos. Parenting X Moderate NE/Low PE [a]	0.012 (0.050)	0.810	-0.065 (0.061)	0.288	0.011 (0.036)	0.753
Pos. Parenting X High NE/High PE [a]	-0.002 (0.058)	0.976	-0.129 (0.066)	0.045	-0.062 (0.042)	0.143

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC1 and 2 are principle components accounting for genetic variation due to race/ethnicity; NE = negative emotional reactivity; PE = positive emotional reactivity.**

[a] The Low NE/Low PE latent profile is the reference group.

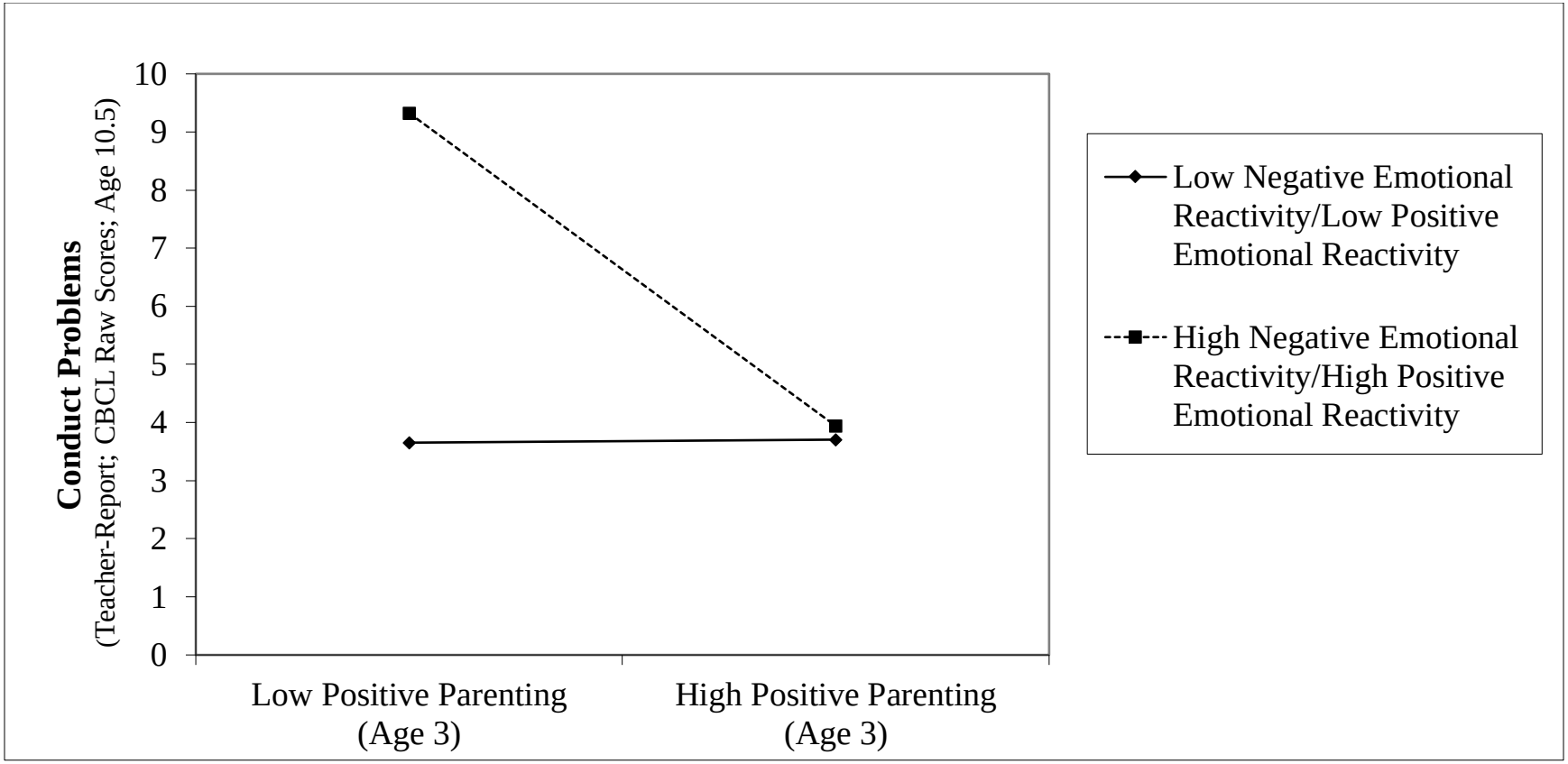

Figure 15. Positive parenting (age 3) x emotional reactivity profiles (age 2) in the whole sample.

3.4.1.2 Negative Parenting X Latent Profiles of Emotional Reactivity (Table 8)

The same models were recomputed replacing age 3 positive parenting with age 3 negative parenting. In contrast to univariate analyses, negative parenting at age 3 was unrelated to age 10.5 CP according to parent-, teacher-, and youth-report. Children's emotional reactivity profiles at age 2 were also unrelated to parent- or youth-reported CP, and there was no evidence of an interaction between age 2 emotional reactivity and age 3 negative parenting in relation to parent- or youth-reported CP. Consistent with models involving positive parenting, emotional reactivity at age 2 predicted teacher-reported CP, such that children in the High NE/High PE profile showed higher levels of CP at age 10.5 compared to youth in the Low NE/Low PE profile. Finally, as predicted by Hypothesis 4, children's emotional reactivity profiles at age 2 interacted with negative parenting at age 3 in predicting teacher-reported CP at age 10.5. As shown in Figure 16, at low levels of age 3 negative parenting, children in the High NE/High PE and Low NE/Low PE profiles did not significantly differ with respect to age 10.5 CP. However, at high levels of age 3 negative parenting, children in the High NE/High PE showed significantly higher levels of teacher-reported CP at age 10.5 compared to children in the Low NE/Low PE profile.

Table 8. Aim 4: Age 2 emotional reactivity profiles by age 3 negative parenting predicting age 10.5 conduct problems (CP)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.142 (0.044)	0.001	0.26 (0.054)	0.000	0.07 (0.053)	0.190
Intervention Status	0.064 (0.049)	0.196	-0.047 (0.055)	0.392	0.027 (0.049)	0.586
Pittsburgh Site	0.070 (0.066)	0.286	0.110 (0.069)	0.115	0.104 (0.055)	0.060
Virginia Site	-0.058 (0.053)	0.273	0.047 (0.074)	0.525	-0.010 (0.054)	0.847
Family Income	-0.043 (0.059)	0.463	-0.108 (0.060)	0.070	-0.096 (0.036)	0.008
Ancestry PC1	0.014 (0.061)	0.814	0.113 (0.076)	0.135	0.086 (0.056)	0.126
Ancestry PC2	0.057 (0.036)	0.110	0.063 (0.034)	0.070	-0.004 (0.035)	0.918
Moderate NE/High PE [a]	-0.036 (0.040)	0.374	0.015 (0.045)	0.743	-0.024 (0.041)	0.558
Moderate NE/Low PE [a]	0.002 (0.053)	0.977	0.002 (0.058)	0.974	-0.008 (0.038)	0.823
High NE/High PE [a]	0.044 (0.052)	0.398	0.186 (0.055)	0.001	0.068 (0.046)	0.136
Neg. Parenting (Age 3)	0.085 (0.053)	0.105	0.094 (0.070)	0.178	0.062 (0.066)	0.341
Neg. Parenting X Moderate NE/High PE [a]	-0.011 (0.031)	0.716	0.026 (0.059)	0.657	0.049 (0.077)	0.520
Neg. Parenting X Moderate NE/Low PE [a]	-0.029 (0.057)	0.611	0.064 (0.063)	0.305	0.000 (0.043)	0.998
Neg. Parenting X High NE/High PE [a]	0.031 (0.055)	0.571	0.152 (0.055)	0.006	0.017 (0.039)	0.670

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC1 and 2 are principle components accounting for genetic variation due to race/ethnicity; NE = negative emotional reactivity; PE = positive emotional reactivity.**

[a] The Low NE/Low PE latent profile is the reference group.

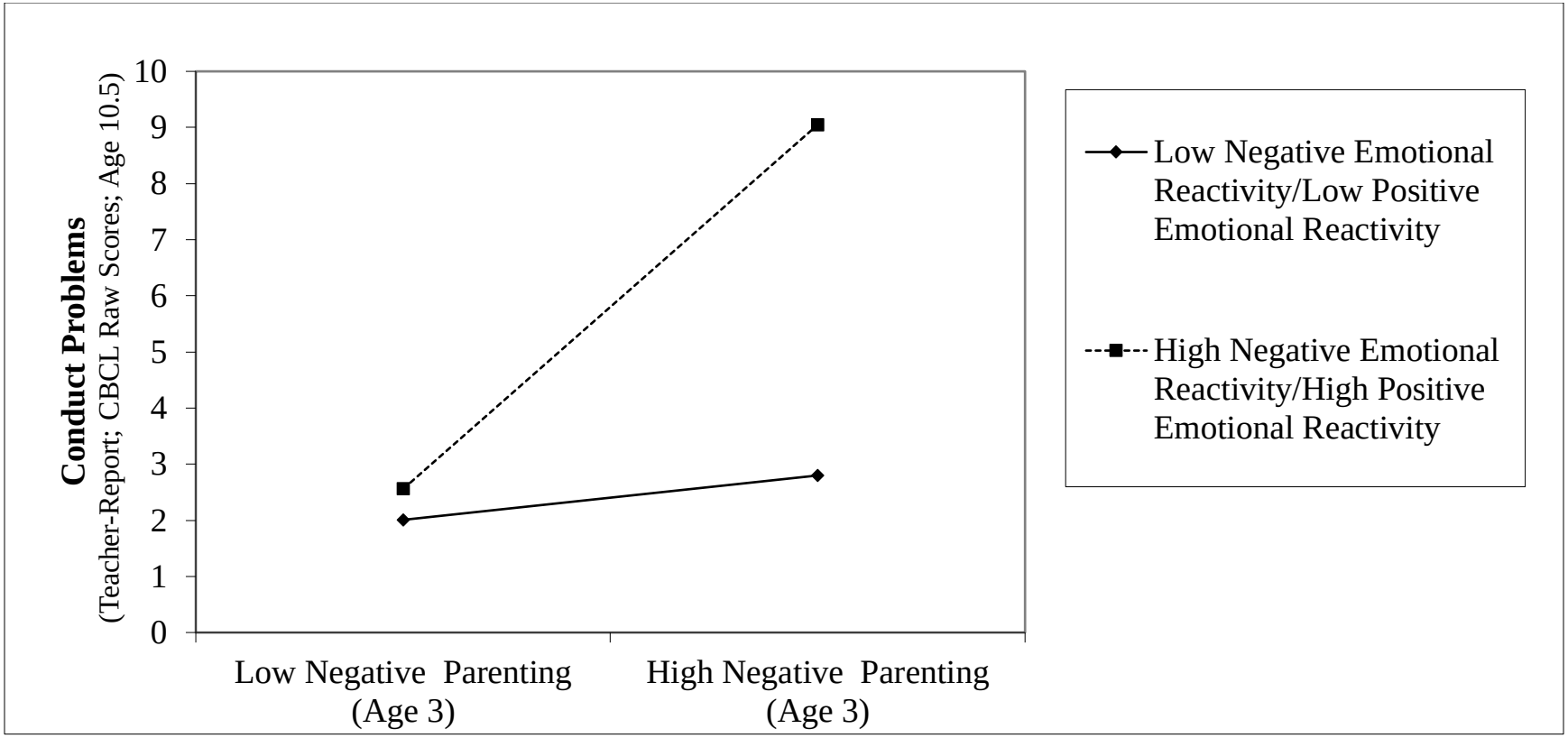

Figure 16. Negative parenting (age 3) x emotional reactivity profiles (age 2) in the whole sample.

Note. **The simple slope for youth in the Low Negative Emotional (NE) Reactivity/Low Positive Emotional (PE) Reactivity latent profile: $t(24) = 1.35$, $p = 0.188$. The simple slope for youth in the High NE Reactivity/High PE Reactivity latent profile: $t(24) = 3.99$, $p = 0.001$. Low NE/Low PE and High NE/Low PE reactivity profiles are significantly different from each other at negative parenting values less than -2.649 and greater than -0.304. The proportion of interaction index (PoI) = 0.99.**

3.4.1.3 Positive and Negative Parenting X Cumulative Susceptibility Scores (Tables 9 and 10)

Our final set of analyses tested interactions between cumulative susceptibility scores and age 3 positive and negative parenting in relation to age 10.5 CP. As previously noted, a cumulative susceptibility score was created for each child based on threshold cut points for observed NE at age 2, parent-reported EC at age 2, and GPS. Children received a score of 0 if their score fell below the indicator's threshold and a score of 1 if their score was above this threshold. For each indicator, we split groups into the bottom ¾ (i.e., score of 0) and top ¼ (i.e., score of 1) of risk[1]. Scores on each indicator were summed for each child, resulting in a single cumulative score that initially ranged from 0 to 3. Descriptive statistics indicated that 38.65% (n = 199) of youth had cumulative susceptibility scores of 0, reporting scores below the threshold cut points for observed NE, parent-reported EC, and GPS. Further, 47.39% (n = 244) had cumulative susceptibility scores of 1, 13.59% (n = 70) had scores of 2, and less than 1% (n = 2) had scores of 3. As there were very few participants who met the threshold cut points on all three indicators, youth with scores of 2 or 3 were combined together, resulting in cumulative susceptibility scores that ranged from 0 to 2.

Analyses computed with the entire sample and Caucasian youth-only indicated that cumulative susceptibility scores were not predictive of parent-, teacher-, or youth-reported CP in middle childhood. However, for non-Caucasian youth, higher cumulative susceptibility scores predicted lower levels of parent-reported CP (significant for positive parenting models [Table S11]

[1] As we had no a priori criteria for establishing thresholds, we initially used multiple cut points, including groups split at the median, those split into the bottom 3/4 and top 1/4, and those split into the bottom 2/3 and top 1/3. Notably, the same pattern of findings emerged regardless of the cut-point used.

and marginal for negative parenting models [Table S13]). Interactions between cumulative susceptibility scores and age 3 positive and negative parenting were non-significant regardless of whether analyses were computed for the entire sample or separately by race.

Table 9. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using aggression genome-wide polygenic scores)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.140 (0.043)[1,2]	0.001	0.249 (0.056)[2]	0.000	0.068 (0.053)	0.196
Intervention Status	0.069 (0.048)[1]	0.145	-0.042 (0.058)	0.473	0.027 (0.049)	0.579
Pittsburgh Site	0.076 (0.06)[2]	0.208	0.119 (0.066)	0.073	0.098 (0.055)	0.071
Virginia Site	-0.062 (0.048)[1]	0.197	0.057 (0.073)	0.433	0.001 (0.057)	0.984
Family Income	-0.020 (0.055)	0.710	-0.063 (0.057)	0.273	-0.080 (0.036)[1, 2]	0.026
Ancestry PC1	0.039 (0.062)	0.532	0.135 (0.074)	0.067	0.091 (0.056)	0.101
Ancestry PC2	0.028 (0.037)	0.460	0.02 (0.048)	0.680	-0.012 (0.037)	0.748
Cumulative Susceptibility Score	-0.085 (0.05)[2]	0.087	0.038 (0.063)	0.550	0.021 (0.041)	0.602
Pos. Parenting (Age 3)	-0.106 (0.050)[1]	0.033	-0.063 (0.067)	0.347	-0.019 (0.048)	0.686
Cumulative Score X Pos. Parenting	0.044 (0.050)	0.376	-0.042 (0.075)	0.578	0.024 (0.042)	0.561

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC1 and 2 are principle components accounting for genetic variation due to race/ethnicity. Model parameters reported separately by race can be found in the Appendix at the end of the document.**

[1] Significant for Caucasian participants. [1] Marginally significant for Caucasian participants. [2] Significant for non-Caucasian participants. [2] Marginally significant for non-Caucasian participants.

Table 10. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using aggression genome-wide polygenic scores)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.131 (0.043)[1,2]	0.002	0.241 (0.056)[2]	0.000	0.066 (0.053)	0.217
Intervention Status	0.064 (0.048)[1]	0.181	-0.041 (0.055)	0.455	0.029 (0.048)	0.554
Pittsburgh Site	0.089 (0.061)[2]	0.143	0.111 (0.065)[2]	0.088	0.094 (0.054)[2]	0.083
Virginia Site	-0.061 (0.049)[1]	0.210	0.032 (0.072)	0.654	-0.005 (0.056)	0.921
Family Income	-0.024 (0.055)	0.665	-0.065 (0.058)	0.265	-0.074 (0.035)[1,2]	0.035
Ancestry PC1	0.040 (0.063)	0.527	0.142 (0.073)	0.053	0.093 (0.056)	0.096
Ancestry PC2	0.034 (0.037)	0.353	0.017 (0.044)	0.696	-0.016 (0.036)	0.657
Cumulative Susceptibility Score	-0.089 (0.050)[2]	0.076	0.022 (0.063)	0.731	0.017 (0.041)	0.684
Neg. Parenting (Age 3)	0.094 (0.050)	0.061	0.135 (0.067)[1,2]	0.044	0.078 (0.054)[2]	0.148
Cumulative Score X Neg. Parenting	-0.058 (0.051)	0.262	0.014 (0.076)	0.855	-0.010 (0.043)	0.810

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC1 and 2 are principle components accounting for genetic variation due to race/ethnicity. Model parameters reported separately by race can be found in the Appendix available at the end of the document.**

[1] **Significant for Caucasian participants.** [1] **Marginally significant for Caucasian participants.** [2] **Significant for non-Caucasian participants.** [2] **Marginally significant for non-Caucasian participants.**

4.0 DISCUSSION

Although differential susceptibility theory has received increasing empirical support in the last decade (Belsky & Pluess, 2009, 2012; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van Ijzendoorn, 2011), several unresolved issues remain. The current study sought to address particularly important questions concerning patterns of convergence or divergence across various markers, time scales, and methods of assessing differential susceptibility. First, we examined the extent to which different markers of susceptibility, specifically phenotypic (i.e., age 2 negative emotionality (NE) and effortful control (EC)) and genotypic characteristics, identify the same or different groups of children as being most susceptible to their caregiving environment during toddlerhood. Second, we sought to identify whether a subset of youth are differentially reactive to their parent's display of positive and negative emotions during moment-to-moment parent-child interactions. We then examined whether being differentially reactive on a short-term time scale generalizes to differential susceptibility across longer periods of development. Finally, although not included as a formal aim of the project, we were interested in examining whether different methods of assessing environmental susceptibility (i.e., parent-report and observational measures of temperament) are capturing the same subgroups of children.

Regarding the issue of associations between early indictors of toddler's observed NE and concurrent parent-reported effortful control EC, as expected, negative associations were evident. However, contrary to our first hypothesis predicting modest associations between children's genome-wide polygenic scores (GPS) and these two temperament dimensions, neither observed NE nor parent-reported EC at age 2 were related to GPS in univariate analyses. Unfortunately, limited variability in GPS in the current study precluded the identification of meaningful subgroups

using this indicator. The inability to include GPS in a Latent Profile Analysis (LPA) with age 2 NE and age 2 EC meant that GPS had to be estimated as an independent continuous predictor to be included in subsequent moderation analyses. Thus, the omission of GPS from the LPA prevented us from formally examining how genotypic and phenotypic characteristics co-occur in the same individuals. To our knowledge, the current study is among the first to include polygenic scores in a LPA. As this analytic technique is a person-oriented approach, it is quite possible that polygenic scores, while potentially useful for population-level prediction, are not yet suitable for making predictions about specific individuals.

Genome-wide polygenic scores (GPS) and age 2 latent profiles of EC and NE were then tested as moderators of the association between age 3 positive and negative parenting and age 10.5 CP. Analyses computed with the entire sample showed direct effects of age 2 EC/NE profiles on age 10.5 CP, with children in the Average EC/High NE profile showing significantly higher levels of teacher-reported CP than those in the Average EC/Low NE profile. Further, profiles of EC/NE at age 2 moderated the effects of positive parenting when predicting teacher-reported CP, such that for youth with Average EC/High NE, lower positive parenting at age 3 predicted greater CP at age 10.5 compared to youth with Average EC/Low NE. For youth with Average EC/Low NE, direct effects of positive parenting on CP in middle childhood were not evident. Thus, person-oriented analyses computed with EC and NE yielded important information about how multiple facets of a child's temperament co-occur and influence one another.

The heightened sensitivity of children with Average EC/High NE to lower positive parenting is consistent with findings from a small but growing body of research exploring interactions between the regulatory and reactive components of temperament (Eisenberg et al., 2000; Moran et al., 2013; Rothbart & Bates, 2006). For example, Moran, Lengua, and Zalewski

(2013) found that children with higher observed frustration *and* lower effortful control at age 3 showed elevated externalizing problems a year later compared to children who also showed high levels of frustration in toddlerhood but greater effortful control. Eisenberg and colleagues (2004) showed a similar pattern of findings in an older sample of youth, with children's dispositional anger at age 6 moderating the effects of concurrently assessed effortful control on externalizing problems at age 8 (Eisenberg et al., 2004). Specifically, effortful control in early childhood was a better predictor of later externalizing problems for children prone to feeling angry than for children who do not experience anger as often or intensely. Thus, findings from the current study and prior research collectively suggest that a predisposition to experience negative emotions without sufficient skills to modulate affective arousal confers increased risk for youth CP and externalizing problems more broadly from toddlerhood (Moran et al., 2013) through early and middle childhood (Eisenberg et al., 2000, 2004). In contrast, children who are prone to experiencing intense negative emotions may be buffered from developing early-onset CP if they are able to regulate their emotional reactivity. By regulating the intensity and duration of their emotional experience, children may be better equipped to identify and execute adaptive responses to emotionally provocative situations and to refrain from less socially acceptable behaviors such as aggression.

Although profiles of EC/NE at age 2 moderated the effects of age 3 positive parenting on teacher-reported CP at age 10.5 in whole-sample *and* race-specific analyses, the specific pattern of interaction effects differed across groups. While the interaction pattern identified in whole sample analyses was also evident when analyses were computed with African American and Latina youth only, the interaction took a different, unexpected form for Caucasian youth. Specifically, for Caucasian youth with Average EC/High NE at age 2, exposure to more positive parenting at age 3 predicted *elevated* CP in middle childhood relative to youth with Average EC and low NE.

Notably, finding replicated across all informants of youth CP (i.e., parent-, teacher-, and youth). One potential explanation for these unexpected findings may relate to the fact we incorporated a unitary construct of youth CP and did not distinguish between different subtypes of antisocial behavior which may show distinct temperament profiles and respond differently to the caregiving environment. Particularly relevant to the current findings is the model proposed by Keenan and Shaw (2003) which outlines separate pathways to reactive and proactive antisocial behavior. The authors argue that the pathway to reactive aggression begins with an irritable and poorly regulated infant who is paired with a parent that is overly responsive to their emotions and places few demands on them. Although responsive parenting is generally linked with positive social and emotional outcomes, these parents may over-compensate for their child's poor self-regulation skills and deprive their child of opportunities to regulate their own emotions. It is possible that our broad measure of positive parenting in the current study was not well-suited for capturing these types of parenting behaviors which may initially appear to be generally positive. Thus, it would beneficial to examine whether the profiles of effortful control (EC) and negative emotionality (NE) identified in the present study are differentially related to various subtypes of CP.

A number of researchers have posited that genetic and temperament characteristics may be identifying the same groups of children through different means. For example, Tung and colleagues (2018) noted, "… perhaps the 'sensitive' individuals separately identified by each trait (e.g., individuals with DRD4 7R allele or high in negative emotionality) are the *same* individuals across multiple markers of susceptibility." Similar speculations also have been made by Belsky and Pluess (2009, 2013) and Slagt and colleagues (2016), as carriers of plasticity alleles for *5-HTTLPR* (Auerbach et al., 2001; Holmboe et al., 2011) and *DRD4* (Holmboe et al., 2011; Ivorra et al., 2011; Oniszczenko & Dragan, 2005) have demonstrated higher levels of negative

emotionality and emotional reactivity in infancy and toddlerhood. However, while these findings suggest that phenotypic and genotypic characteristics may be identifying the same "susceptible" individuals through different means, this important issue can only truly be tested by examining multiple markers of susceptibility within the *same* children. To our knowledge, the current study represents the first attempt to directly test this issue empirically by using LPA, a person-oriented approach to examining how continuously measured characteristics co-occur within individuals. However, limited variability in the aggression GPS precluded us from including these scores in the LPA, an issue that also occurred when using a polygenic score representing environmental sensitivity based on identical twin differences (ages 5-18 years; M_{age} = 11.28) in emotional problems (Keers et al., 2016). Nonetheless, analyses using cumulative susceptibility scores revealed that less than 1% of the current sample was above the threshold cut off for observed NE at age 2, parent-reported EC at age 2, *and* GPS. As the majority of youth showed heightened susceptibility based only on one or two markers, these findings provide initial evidence suggesting divergence across different markers of environmental susceptibility. However, it is important to acknowledge that cumulative susceptibility scores were created using threshold cut-offs that dichotomized each susceptibility factor into scores of zero and one. While this approach yielded important findings, it likely also resulted in the loss of important information about co-occurring patterns of different susceptibility markers. Thus, further research is needed that employs LPA which is ideally suited for assessing patterns of convergence and divergence across *continuously-measured* markers of susceptibility.

Our inability to include the GPS in the LPA because of limited variability speaks to the need for richer methodological techniques for assessing genetic phenomena. The current study employed a data-driven approach in which SNPs previously found to be associated with an

outcome of interest in a GWAS are combined into a polygenic score. Although this method leverages the power of GWAS to detect small effects across the genome, allowing for polygenic scores composed of hundreds or thousands of SNPs, it is also blind to theory and vulnerable to statistical artifacts (Belsky & Israel, 2014). An alternative approach to be considered in future research involves forming polygenic scores from a small number of candidate SNPs shown to be relevant to a trait or behavior. While this approach to creating polygenic scores is more theoretically-informed than the GWAS approach, neither approach integrates information about functional biological relevance in their selection of SNPs. Gene set enrichment analysis (GSEA) offers an attractive alternative to creating polygenic scores, using bioinformatics to model functional genetic networks, improve the measurement of genetic susceptibility, and reduce the number of statistical tests performed (Mooney & Wilmot, 2015). It remains to be seen whether these more biologically-informed polygenic scores are better suited for inclusion in LPA.

As previously noted, findings indicated moderate negative associations between parent-reported effortful control and observed negative emotionality at age 2. Unfortunately, however, questionnaire and observational measures of the same temperament construct were not available. Thus, the current study cannot address whether different methods of assessing the same susceptibility factor are capturing the same subgroups of children. This issue remains an extremely important area for future research, especially as the field grapples with how to best translate empirically-supported intervention and prevention programs for youth problem behaviors to clinical and educational settings in ways that minimize time and costs while ensuring adherence to evidence-based practices.

Turning to Aim 2, as expected, child negative emotional (NE) reactivity at age 2 was moderately negatively correlated with positive emotional (PE) reactivity concurrently. With

respect to our question concerning children who show similar profiles of emotional reactivity during observed parent-child interactions, LPA identified four profiles of youth varying in levels of both dimensions. As expected, the High NE/High PE reactivity profile included only a small minority of children (5.2%) who are emotionally reactive "for better and for worse," responding with positive emotions to their parents' positive emotions and with negative emotions to their parents' display of negative emotions.

A key assumption of the differential susceptibility model is that the *same* children who are most adversely affected by negative environments may also disproportionately benefit from positive environments. However, due in part to the practical and ethical challenges of experimentally exposing children to both positive and negative caregiving environments, few studies have been able to directly test this assumption. The current study sought to advance research on differential susceptibility by utilizing observations of parent-child interactions to examine variation in each child's exposure to both positive and negative parenting in toddlerhood. Thus, while past studies have typically employed between-person designs and established relations at the level of the group or population, the current approach to measuring children's positive and negative emotional reactivity yielded a within-person design capable of evaluating an important but often untested assumption of differential susceptibility theory.

Study findings showed that some children are emotionally reactive "for better and for worse" (i.e., High NE/High PE), while others are generally emotionally unreactive (i.e., Low NE/Low PE) to their caregiving environment. However, it is important to note that levels of NE and PE reactivity did not converge for children in the two other profiles of emotional reactivity identified via LPA. Although both profiles showed moderate levels of NE reactivity, children in one profile showed high levels of PE reactivity (i.e., Moderate NE/High PE), while those in the

other showed low PE reactivity. These findings corroborate prior literature showing that negative and positive emotionality are distinct, albeit moderately correlated, constructs in infants and children and that high levels of negative emotional (NE) reactivity are not uniformly associated with low levels of positive emotional (PE) reactivity within individuals (Belsky, Hsieh, & Crnic, 1996). For example, Belsky and colleagues (1996) found that in a sample of 12-13-month old infants, a two-construct model of PE and NE reactivity fit the data better than a one-construct model. The findings from these studies provide strong support for a conceptualization of emotional reactivity that distinguishes between positive and negative emotionality rather than assuming them to be opposite ends of a single continuum. However, from a methodological standpoint, it is noteworthy that positive and negative emotional reactivity were assessed during the same tasks in the present study, which may account for some of the interdependence between these constructs. Future research would benefit from using separate tasks to better differentiate the level of independence or dependence between negative and positive emotionality.

After identifying profiles of emotional reactivity at age 2, we then sought to examine the extent to which children's NE and PE reactivity translates to differential susceptibility across longer periods of development. We found partial support for the hypothesis that children who react more strongly to their parents' display of emotions during moment-to-moment interactions show stronger longitudinal associations between negative and positive parenting in toddlerhood in relation to age 10.5 conduct problems (CP). As expected, compared to those in the Low NE/Low PE profile, children in the High NE/High PE profile showed higher levels of teacher-reported CP at age 10.5 when exposed to low levels of positive parenting at age 3 or high levels of negative parenting at age 3. However, overall results did not support a pattern of differential susceptibility, as youth with High NE and High PE reactivity did not show heightened sensitivity to more positive

or less negative parenting behaviors at age 3. Youth in the Low NE/Low PE profile showed similar levels of age 10.5 CP regardless of whether they were exposed to more or less positive or negative parenting at age 3.

Although studies increasingly support negative emotionality as a marker of differential susceptibility (Slagt et al., 2016), prior work in this area has primarily focused on long-term developmental changes in response to naturally occurring variation in the environment or interventions. Findings from the current study suggest that children who are differentially reactive to their parents' display of emotions during observed interactions coded at a molecular level may be more susceptible to adverse but not supportive caregiving environments across longer periods of development. In fact, the only other study to empirically examine this issue found that in a sample of children ages 4 to 6, longitudinal associations of positive and negative parenting with externalizing and prosocial behaviors assessed a year later were similar for children in the high emotional reactivity and average emotional reactivity profiles (Slagt et al., 2019). Although the current study also found no evidence that emotional reactivity profiles at age 2 were differentially susceptible to more positive caregiving environments at age 3, as previously noted, children high in PE and NE reactivity at age 2 *were* more susceptible to harsh caregiving at age 3. This discrepancy in findings between the two studies emphasize that further work in this area is needed before we can conclude whether emotional reactivity coded molecularly via observation does or does not translate to differential susceptibility over many years.

Further, we would be remiss if we did not acknowledge that age 10.5 CP was the sole indicator of child adjustment in middle childhood. Simply treating the absence of CP as the positive end of psychological functioning likely limited our ability to adequately differentiate between models of environmental sensitivity, such as differential susceptibility, diathesis stress, and

vantage sensitivity. Thus, although the interaction between age 2 emotional reactivity and age 3 parenting was consistent with a diathesis stress framework, future replication studies that incorporate the full range of environments and outcomes, from positive to negative, are needed to further clarify the relationship between short-term reactivity and long-term differential susceptibility.

It is important to note that interactions between age 2 emotional reactivity profiles and age 3 parenting were largely specific to teacher-rated CP at aged 10.5 and were not predictive of parent- or youth-rated CP at this age. Cross-informant discrepancies in the assessment of youth CP are more often "the rule, rather than the exception" (De Los Reyes et al., 2015; Ferdinand et al., 2004) and likely reflect varying perspectives of a child's behavior and actual differences in a child's display of CP across various contexts. These results reinforce the value of examining youth CP across multiple informants and contexts when assessing patterns of environmental sensitivity.

4.1 LIMITATIONS

The current study should be interpreted within the context of several limitations. First, the current sample was originally recruited from WIC nutrition supplement centers and further screened and deemed eligible based on the presence of additional socioeconomic (i.e., parental educational attainment), family (maternal depression, substance use) and/or child risk (i.e., high levels of CP); thus, findings are limited in generalizability to high-risk male and female children and families from low-SES backgrounds living in rural, urban, and suburban communities.

Second, children's emotional reactivity was operationalized to reflect the percentage of interaction time that children changed their emotions *based on* prior changes in their parents'

emotions. Thus, the maximum amount of emotional reactivity that a child could display was limited by the frequency with which a parent changed their emotions. For example, if the *child* frequently alternated between displays of positive and negative affect throughout the observation tasks but their parent was consistently positive, then such a child would have likely been coded as low on PE and NE reactivity. Although this coding approach was intentional because of an interest in examining children's affective responses to *specific* environmental stimuli (i.e., their differential reactivity to their parents' change in emotional expressions), defining children's emotional reactivity in this way may have limited our ability to capture children's "true" levels of PE and NE reactivity.

Third, although we focused specifically on children's emotional reactivity to their parents' display of positive and negative emotions, reactivity to environmental stimuli also encompasses a child's behavioral response. It remains to be seen whether profiles of emotional reactivity identified on the basis of verbal and physical codes from the Relationship Affect Coding System converge with those identified in the present study based on affect codes alone.

Fourth, the analytic sample (N = 515) was relatively modest compared to large-scale epidemiological investigations. Notably, a smaller sample permitted longitudinal measurement spanning approximately ten years, the rigorous assessment of parenting using observational methods, and a multi-informant approach to assessing youth CP. Although these advantages have posed longstanding challenges in the genetic epidemiology of complex phenotypes, the small sample size may have limited our ability to detect and sufficiently probe all interaction effects.

Finally, it is critical to acknowledge GPS were derived from GWAS' of European children, suggesting that they may not adequately capture genetic risk for those of non-European descent in our sample. To address concerns of population stratification, models were first computed with the

entire sample controlling for ancestry principal components and were then followed by analyses computed separately for Caucasian and non-Caucasian youth. Although both approaches are certainly preferred over models that make no attempt to control for population stratification, it is important to remember that these approaches assume that the underlying structure of the GPS is the same across racial groups (i.e., the GPS are computed in the same way for those of European and non-European descent). This issue is highly problematic, as research has shown that PRS derived from GWAS of European-ancestry samples can misestimate risk when applied to non-European cohorts (Kim et al., 2018). However, this concern is not a problem specific to this study, but rather a problem for the field of behavioral genetics more generally. While this historically Eurocentric approach to genetic research may not seem to be much of a concern at face value, the magnitude of this issue becomes more apparent as the prospect of using polygenic scores to guide clinical care gains traction (De La Vega & Bustamante, 2018). Although progress in this area is exciting and has the potential to lead to improved diagnostic sensitivity and more individualized intervention and prevention approaches, it also has the potential to further exacerbate existing racial and ethnic disparities in receiving quality health care. Thus, the inclusion of more diverse populations in GWAS and biobanking needs to be prioritized so that all can benefit from the progress being made in genomic medicine (Popejoy & Fullerton, 2016).

4.2 IMPLICATIONS AND FUTURE DIRECTIONS

Whether phenotypic (i.e., temperament) and genotypic characteristics converge or diverge in identifying youth most susceptible to their environment remains an important question for the field. While the present study sought to address this gap in the field, methodological limitations

(i.e., restricted variability in the aggression GPS) prohibited us from employing a person-oriented approach to identify patterns in how genetic characteristics and multiple facets of temperament (i.e., age 2 EC and NE) co-occur across individuals. While the current study was unable to shed light on this important issue, it constitutes no more than an initial attempt to answering this fundamental question. Future work on this topic is needed before research on differential susceptibility can be translated to clinical practice, permitting the systematic screening of young children based on susceptibility markers. Specifically, if phenotypic and genotypic indicators diverge in identifying the most susceptible youth, the findings would suggest that different markers are identifying different subgroups of youth. Thus, although some susceptible youth may be captured by genetic markers, another subgroup of children most susceptible to their environment and later maladaptive outcomes may be better identified based on their levels of negative emotionality or emotional reactivity. This divergence would underscore the importance of developing screening procedures that incorporate multiple markers of susceptibility to identify children most likely to benefit from intervention. Thus, based on the potential implications of the differential susceptibility perspective for assessment, prevention, and intervention services, future studies examining how multiple markers of susceptibility co-occur within individuals are needed.

Despite making significant contributions to the literature, the present study does not elucidate mechanisms of differential influence. This area of work is largely speculative and under-developed, and future work is needed to identify potential neural and cognitive mechanisms that may explain how reverse allelic associations may occur across adverse and advantageous environments. The identification of these specific processes may lead to more cost-effective and robust intervention and prevention efforts for youth CP.

APPENDIX A SUPPLEMENTAL TABLES AND FIGURES

Table S1. Latent profile analysis with effortful control, negative emotionality, aggression genome-wide polygenic scores

# Profiles	BIC	Adjusted BIC	AIC	Loglikelihood	Entropy	Profile Sizes (%)
1	1998.763	1979.720	1973.796	-980.898	NA[a]	100
2	1666.064	1634.326	1624.452	-802.226	0.960	89.93/10.07
3	1498.491	1454.057	1440.234	-706.117	0.971	86.75/8.94/4.30
4	1405.391	1348.262	1330.489	-647.245	0.946	75.12/13.53/7.18/4.17

***Note*. AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; BLRT = Bootstrap Likelihood Ratio Test for k versus k-1 classes;**

[a] Entropy is not available for one-class models.

Table S2. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by MZ twin difference polygenic scores and profiles of effortful control and negative emotionality

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.041 (0.048)	0.397	0.081 (0.055)[2]	0.142	0.018 (0.052)	0.731
Child Age	-0.058 (0.046)	0.208	-0.007 (0.057)[1]	0.896	0.071 (0.049)	0.145
Intervention Status	0.087 (0.048)	0.072	0.014 (0.055)	0.797	0.030 (0.052)	0.560
Pittsburgh Site	0.027 (0.064)	0.674	0.068 (0.068)[2]	0.319	0.140 (0.057)[2]	0.013
Virginia Site	-0.107 (0.052)	0.042	0.014 (0.066)	0.831	0.035 (0.060)	0.552
Family Income	-0.051 (0.055)	0.355	-0.099 (0.067)	0.139	-0.090 (0.038)	0.017
Ancestry PC1	-0.009 (0.067)	0.898	0.127 (0.079)	0.110	0.079 (0.059)	0.180
Ancestry PC2	0.048 (0.044)	0.280	0.018 (0.048)	0.710	0.034 (0.036)	0.347
Class 1 (Age 2)	0.028 (0.069)	0.679	-0.202 (0.128)[2]	0.114	-0.071 (0.076)	0.352
Class 2 (Age 2)	0.102 (0.079)[2]	0.195	-0.125 (0.114)[2]	0.274	0.061 (0.087)	0.484
GPS	-0.048 (0.049)	0.335	-0.039 (0.065)	0.547	-0.014 (0.041)	0.738
Pos. Parenting (Age 3)	0.183 (0.155)[1]	0.238	-0.382 (0.327)[1,2]	0.242	-0.150 (0.162)[1,2]	0.355
Pos. Parenting X Class 1	-0.294 (0.141)[1]	0.037	0.303 (0.313)[1,2]	0.334	0.110 (0.161)[1]	0.494
Pos. Parenting X Class 2	-0.058 (0.081)[1]	0.471	0.007 (0.149)[1,2]	0.965	-0.027 (0.092)[1]	0.766
GPS X Pos. Parenting	-0.052 (0.043)	0.226	-0.035 (0.069)[1]	0.609	-0.034 (0.037)	0.368

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity; GPS = genome-wide polygenic scores.**

[1] Significant for Caucasian participants. [2] Significant for non-Caucasian participants. [2] Marginally significant for non-Caucasian participants.

Table S3. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by MZ twin difference polygenic scores and profiles of effortful control and negative emotionality

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est.(SE)	P	Est. (SE)	P
Child Sex	0.048 (0.048)	0.312	0.082 (0.061)	0.178	0.025 (0.053)	0.646
Child Age	-0.073 (0.044)	0.103	-0.041 (0.053)	0.441	0.042 (0.048)	0.381
Intervention Status	0.085 (0.050)[1]	0.088	0.011 (0.057)	0.851	0.032 (0.050)	0.517
Pittsburgh Site	0.051 (0.066)	0.435	0.067 (0.069)	0.328	0.152 (0.056)[2]	0.007
Virginia Site	-0.107 (0.054)[1]	0.048	-0.007 (0.072)	0.918	0.035 (0.060)	0.565
Family Income	-0.049 (0.056)	0.379	-0.105 (0.066)	0.113	-0.088 (0.038)	0.020
Ancestry PC1	-0.008 (0.069)	0.903	0.136 (0.082)	0.098	0.068 (0.057)	0.236
Ancestry PC2	0.044 (0.044)	0.316	0.014 (0.049)	0.772	0.028 (0.036)	0.445
Class 1 (Age 2)	0.016 (0.071)	0.820	-0.152 (0.145)	0.293	-0.052 (0.070)	0.460
Class 2 (Age 2)	0.085 (0.081)	0.290	-0.113 (0.122)	0.358	0.062 (0.075)	0.409
GPS	-0.048 (0.050)	0.336	-0.041 (0.070)	0.552	-0.016 (0.042)	0.696
Neg. Parenting (Age 3)	-0.171 (0.177)	0.335	-0.158 (0.410)	0.699	0.100 (0.240)[1]	0.677
Neg. Parenting X Class 1	0.202 (0.166)	0.224	0.259 (0.357)	0.467	-0.065 (0.235)	0.782
Neg. Parenting X Class 2	0.147 (0.072)	0.042	0.245 (0.102)[2]	0.016	0.147 (0.118)[1]	0.210
Polygenic X Neg. Parenting	0.081 (0.041)[2]	0.049	0.079 (0.067)	0.238	-0.002 (0.051)	0.968

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity; GPS = genome-wide polygenic scores.**

[1] Significant for Caucasian participants. [2] Significant for non-Caucasian participants.

Table S4. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using MZ twin difference scores)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.034 (0.047)	0.471	0.089 (0.060)[2]	0.135	0.071 (0.052)	0.177
Intervention Status	0.069 (0.048)[1]	0.152	-0.039 (0.059)	0.511	0.026 (0.049)	0.599
Pittsburgh Site	0.063 (0.062)	0.309	0.092 (0.071)	0.191	0.095 (0.054)[2]	0.079
Virginia Site	-0.068 (0.050)[1]	0.179	0.027 (0.072)	0.709	0.001 (0.056)	0.989
Family Income	-0.036 (0.053)	0.499	-0.088 (0.060)[2]	0.140	-0.073 (0.036)[1, 2]	0.040
Ancestry PC1	0.007 (0.061)	0.907	0.166 (0.074)	0.026	0.108 (0.057)	0.058
Ancestry PC2	0.044 (0.042)	0.288	0.005 (0.054)	0.926	-0.012 (0.038)	0.758
Cumulative Susceptibility Score	-0.072 (0.047)[2]	0.129	0.031 (0.063)	0.624	0.053 (0.043)	0.218
Pos. Parenting (Age 3)	-0.119 (0.050)[1]	0.018	-0.090 (0.069) [2]	0.195	-0.020 (0.048)	0.681
Cumulative Score X Pos. Parenting	0.035 (0.045)	0.440	-0.069 (0.074)	0.351	-0.027 (0.036)	0.453

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity.**

[1] Significant for Caucasian participants. [1] Marginally significant for Caucasian participants. [2] Significant for non-Caucasian participants. [2] Marginally significant for non-Caucasian participants.

Table S5. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using MZ twin difference scores)

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.025 (0.047)	0.595	0.082 (0.059)[2]	0.165	0.068 (0.053)	0.202
Intervention Status	0.061 (0.048)	0.203	-0.045 (0.057)	0.427	0.024 (0.048)	0.614
Pittsburgh Site	0.075 (0.062)[2]	0.232	0.088 (0.07)[2]	0.208	0.088 (0.054) [2]	0.105
Virginia Site	-0.071 (0.052)[1]	0.172	-0.001 (0.074)	0.989	-0.018 (0.054)	0.739
Family Income	-0.037 (0.053)	0.479	-0.088 (0.061)[2]	0.149	-0.065 (0.036)[1, 2]	0.066
Ancestry PC1	0.013 (0.061)	0.835	0.173 (0.075)	0.021	0.117 (0.056)	0.038
Ancestry PC2	0.056 (0.042)	0.177	0.006 (0.050)	0.908	-0.014 (0.038)	0.717
Cumulative Susceptibility Score	-0.073 (0.048)[2]	0.122	0.035 (0.064)	0.585	0.053 (0.043)	0.213
Neg. Parenting (Age 3)	0.094 (0.050)	0.062	0.144 (0.061)[2]	0.019	0.062 (0.055)	0.257
Cumulative Score X Neg. Parenting	0.019 (0.048)	0.699	0.089 (0.080)	0.267	0.085 (0.045)[1]	0.057

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity.**

[1] Significant for Caucasian participants. [1] Marginally significant for Caucasian participants.

[2] Significant for non-Caucasian participants. [2] Marginally significant for non-Caucasian participants.

Table S6. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by aggression genome-wide polygenic scores (GPS) and age 2 latent profiles of effortful control (EC) and negative emotionality (NE) – Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.142 (0.055)	0.009	0.061 (0.036)	0.088	0.014 (0.043)	0.750
Child Age	-0.106 (0.053)	0.045	-0.076 (0.038)	0.045	0.020 (0.042)	0.638
Intervention Status	0.131 (0.067)	0.050	0.005 (0.035)	0.891	0.035 (0.049)	0.469
Pittsburgh Site	-0.015 (0.065)	0.819	0.004 (0.032)	0.896	0.061 (0.044)	0.168
Virginia Site	-0.098 (0.048)	0.039	0.036 (0.041)	0.376	0.057 (0.057)	0.318
Family Income	-0.026 (0.061)	0.671	-0.049 (0.029)	0.096	-0.054 (0.036)	0.127
Avg. EC/Low NE (Age 2)[a]	-0.057 (0.044)	0.195	-0.081 (0.057)	0.156	-0.020 (0.047)	0.670
Avg. EC/Moderate NE (Age 2)[a]	-0.048 (0.083)	0.562	-0.076 (0.055)	0.168	-0.019 (0.047)	0.681
GPS	-0.009 (0.046)	0.854	-0.026 (0.032)	0.423	-0.023 (0.033)	0.491
Pos. Parenting (Age 3)	0.402 (0.131)	0.002	0.653 (0.039)	0.000	0.614 (0.041)	0.000
Pos. Parenting X Avg. EC/Low NE	-0.488 (0.113)	0.000	-0.621 (0.041)	0.000	-0.588 (0.053)	0.000
Pos. Parenting X Avg. EC/Moderate NE	-0.209 (0.061)	0.001	-0.258 (0.050)	0.000	-0.231 (0.047)	0.000
GPS X Pos. Parenting	-0.022 (0.042)	0.600	0.016 (0.025)	0.515	0.002 (0.027)	0.929

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; GPS = genome-wide polygenic scores.**

[a] **The Avg. EC/High NE latent profile is the reference group.**

Table S7. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by aggression genome-wide polygenic scores (GPS) and Age 2 latent profiles of effortful control (EC) and negative emotionality (NE) – Non-Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.105 (0.067)	0.115	0.141 (0.042)	0.001	-0.002 (0.064)	0.972
Child Age	-0.035 (0.063)	0.578	0.037 (0.038)	0.327	0.065 (0.050)	0.192
Intervention Status	0.024 (0.066)	0.710	0.001 (0.034)	0.970	0.022 (0.059)	0.709
Pittsburgh Site	0.148 (0.092)	0.109	0.097 (0.039)	0.013	0.232 (0.074)	0.002
Virginia Site	-0.044 (0.080)	0.580	0.033 (0.038)	0.386	0.056 (0.068)	0.411
Family Income	-0.025 (0.073)	0.730	-0.050 (0.040)	0.217	-0.091 (0.042)	0.029
Avg. EC/Low NE (Age 2)[a]	0.050 (0.095)	0.600	-0.181 (0.052)	0.001	-0.081 (0.076)	0.288
Avg. EC/Moderate NE (Age 2)[a]	0.215 (0.103)	0.037	-0.115 (0.052)	0.028	0.098 (0.095)	0.302
GPS	-0.133 (0.079)	0.093	-0.044 (0.031)	0.164	-0.159 (0.071)	0.026
Pos. Parenting (Age 3)	0.095 (0.158)	0.549	-0.677 (0.033)	0.000	-0.308 (0.155)	0.047
Pos. Parenting X Avg. EC/Low NE	-0.186 (0.158)	0.241	0.571 (0.052)	0.000	0.249 (0.165)	0.132
Pos. Parenting X Avg. EC/Moderate NE	0.083 (0.112)	0.458	0.195 (0.042)	0.000	0.073 (0.088)	0.403
GPS X Pos. Parenting	0.011 (0.076)	0.881	0.008 (0.036)	0.832	0.111 (0.076)	0.144

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; GPS = genome-wide polygenic scores.**

[a] **The Avg. EC/High NE latent profile is the reference group.**

Table S8. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by aggression genome-wide polygenic scores (GPS) and age 2 latent profiles of effortful control (EC) and negative emotionality (NE) – Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.158 (0.061)	0.010	0.098 (0.157)	0.532	0.005 (0.071)	0.939
Child Age	-0.125 (0.066)	0.060	-0.182 (0.281)	0.517	0.008 (0.069)	0.910
Intervention Status	0.169 (0.070)	0.015	0.037 (0.086)	0.670	0.071 (0.078)	0.359
Pittsburgh Site	0.007 (0.088)	0.939	-0.011 (0.066)	0.869	0.126 (0.076)	0.098
Virginia Site	-0.157 (0.060)	0.009	-0.022 (0.089)	0.805	0.071 (0.088)	0.421
Family Income	-0.060 (0.082)	0.468	-0.094 (0.151)	0.533	-0.096 (0.061)	0.117
Avg. EC/Low NE (Age 2)[a]	-0.058 (0.073)	0.422	-0.041 (0.379)	0.915	-0.075 (0.114)	0.510
Avg. EC/Moderate NE (Age 2)[a]	-0.057 (0.110)	0.606	-0.034 (0.343)	0.921	-0.065 (0.105)	0.535
GPS	-0.015 (0.071)	0.831	-0.062 (0.103)	0.544	-0.047 (0.055)	0.394
Neg. Parenting (Age 3)	0.057 (0.237)	0.809	0.555 (0.665)	0.404	-0.396 (0.203)	0.051
Neg. Parenting X Avg. EC/Low NE	0.042 (0.225)	0.851	-0.417 (0.776)	0.591	0.370 (0.207)	0.074
Neg. Parenting X Avg. EC/Moderate NE	0.016 (0.108)	0.880	-0.085 (0.352)	0.809	0.222 (0.067)	0.001
GPS X Neg. Parenting	0.044 (0.063)	0.492	-0.111 (0.168)	0.506	-0.048 (0.055)	0.381

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; GPS = genome-wide polygenic scores.**

[a] The Avg. EC/High NE latent profile is the reference group.

Table S9. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by aggression genome-wide polygenic scores (GPS) and age 2 latent profiles of effortful control (EC) and negative emotionality (NE) – Non-Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.114 (0.066)	0.085	0.238 (0.109)	0.029	0.004 (0.070)	0.949
Child Age	-0.054 (0.055)	0.325	0.011 (0.056)	0.845	0.038 (0.050)	0.447
Intervention Status	0.028 (0.062)	0.655	-0.010 (0.056)	0.859	0.025 (0.060)	0.682
Pittsburgh Site	0.134 (0.082)	0.104	0.151 (0.078)	0.052	0.242 (0.072)	0.001
Virginia Site	-0.043 (0.074)	0.562	0.052 (0.065)	0.424	0.054 (0.066)	0.414
Family Income	-0.032 (0.065)	0.625	-0.071 (0.076)	0.356	-0.101 (0.047)	0.032
Avg. EC/Low NE (Age 2)[a]	0.048 (0.082)	0.556	-0.056 (0.146)	0.701	-0.057 (0.075)	0.449
Avg. EC/Moderate NE (Age 2)[a]	0.174 (0.101)	0.085	-0.019 (0.115)	0.868	0.093 (0.084)	0.272
GPS	-0.094 (0.073)	0.198	-0.038 (0.059)	0.520	-0.143 (0.058)	0.013
Neg. Parenting (Age 3)	-0.238 (0.187)	0.204	-0.396 (0.280)	0.158	0.203 (0.235)	0.389
Neg. Parenting X Avg. EC/Low NE	0.223 (0.178)	0.211	0.463 (0.221)	0.036	-0.162 (0.236)	0.494
Neg. Parenting X Avg. EC/Moderate NE	0.122 (0.083)	0.144	0.296 (0.077)	0.000	0.044 (0.120)	0.711
GPS X Neg. Parenting	-0.125 (0.073)	0.086	0.011 (0.060)	0.860	-0.172 (0.089)	0.052

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention; GPS = genome-wide polygenic scores.**

[a] **The Avg. EC/High NE latent profile is the reference group.**

Table S10. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using aggression genome-wide polygenic scores) - Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.161 (0.056)	0.004	0.146 (0.089)	0.101	0.041 (0.079)	0.605
Intervention Status	0.134 (0.065)	0.041	-0.032 (0.088)	0.720	0.082 (0.084)	0.328
Pittsburgh Site	0.036 (0.079)	0.650	0.128 (0.089)	0.149	0.107 (0.083)	0.198
Virginia Site	-0.090 (0.047)	0.056	0.133 (0.116)	0.253	0.056 (0.100)	0.577
Family Income	-0.039 (0.074)	0.595	-0.071 (0.070)	0.308	-0.111 (0.057)	0.053
Cumulative Susceptibility Score	-0.056 (0.085)	0.513	-0.031 (0.083)	0.704	-0.020 (0.081)	0.806
Pos. Parenting (Age 3)	-0.157 (0.057)	0.006	-0.016 (0.079)	0.842	0.033 (0.070)	0.633
Cumulative Score X Pos. Parenting	-0.089 (0.060)	0.140	-0.090 (0.082)	0.272	0.041 (0.065)	0.531

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention.**

Table S11. Aim 3: Age 3 positive parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using aggression genome-wide polygenic scores) – Non-Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.118 (0.062)	0.055	0.340 (0.073)	0.000	0.085 (0.072)	0.240
Intervention Status	0.027 (0.067)	0.686	-0.048 (0.076)	0.525	0.000 (0.065)	0.996
Pittsburgh Site	0.141 (0.081)	0.081	0.140 (0.091)	0.122	0.123 (0.077)	0.111
Virginia Site	-0.030 (0.076)	0.691	0.027 (0.094)	0.776	-0.001 (0.076)	0.994
Family Income	0.012 (0.072)	0.870	-0.116 (0.083)	0.164	-0.076 (0.044)	0.087
Cumulative Susceptibility Score	-0.118 (0.057)	0.039	0.076 (0.082)	0.352	0.048 (0.048)	0.317
Pos. Parenting (Age 3)	-0.084 (0.071)	0.237	-0.150 (0.095)	0.117	-0.052 (0.061)	0.397
Cumulative Score X Pos. Parenting	0.106 (0.067)	0.111	-0.009 (0.111)	0.938	0.021 (0.052)	0.689

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention.**

Table S12. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores (Using aggression genome-wide polygenic scores) – Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.148 (0.057)	0.009	0.138 (0.088)	0.117	0.044 (0.078)	0.572
Intervention Status	0.119 (0.067)	0.076	-0.016 (0.084)	0.845	0.087 (0.076)	0.255
Pittsburgh Site	0.047 (0.082)	0.568	0.107 (0.084)	0.205	0.104 (0.083)	0.21
Virginia Site	-0.106 (0.054)	0.052	0.094 (0.120)	0.432	0.056 (0.097)	0.561
Family Income	-0.055 (0.074)	0.455	-0.066 (0.073)	0.369	-0.105 (0.053)	0.05
Cumulative Susceptibility Score	-0.066 (0.093)	0.475	-0.064 (0.085)	0.451	-0.022 (0.080)	0.783
Neg. Parenting (Age 3)	0.110 (0.077)	0.150	0.159 (0.096)	0.098	-0.011 (0.084)	0.894
Cumulative Score X Neg. Parenting	-0.055 (0.082)	0.506	-0.055 (0.099)	0.577	0.004 (0.081)	0.956

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention.**

Table S13. Aim 3: Age 3 negative parenting predicting age 10.5 conduct problems (CP) – Moderation by cumulative susceptibility scores – Non-Caucasian youth only

	Parent-Report CP		Teacher-Report CP		Youth-Report CP	
	Est. (SE)	P	Est. (SE)	P	Est. (SE)	P
Child Sex	0.113 (0.062)	0.069	0.328 (0.073)	0.000	0.078 (0.073)	0.288
Intervention Status	0.022 (0.067)	0.746	-0.057 (0.073)	0.436	-0.002 (0.065)	0.976
Pittsburgh Site	0.178 (0.078)	0.023	0.179 (0.078)	0.021	0.13 (0.073)	0.076
Virginia Site	-0.003 (0.075)	0.971	0.039 (0.088)	0.661	0.001 (0.070)	0.990
Family Income	0.019 (0.071)	0.786	-0.107 (0.084)	0.201	-0.078 (0.045)	0.086
Cumulative Susceptibility Score	-0.109 (0.057)	0.058	0.064 (0.078)	0.412	0.046 (0.049)	0.349
Neg. Parenting (Age 3)	0.087 (0.067)	0.193	0.140 (0.085)	0.100	0.119 (0.069)	0.083
Cumulative Score X Neg. Parenting	-0.052 (0.068)	0.445	0.033 (0.094)	0.727	-0.021 (0.056)	0.712

Note. **Child Sex: 0 = Female; 1 = Male; Intervention: 0 = Control, 1 = Family Check-Up Intervention.**

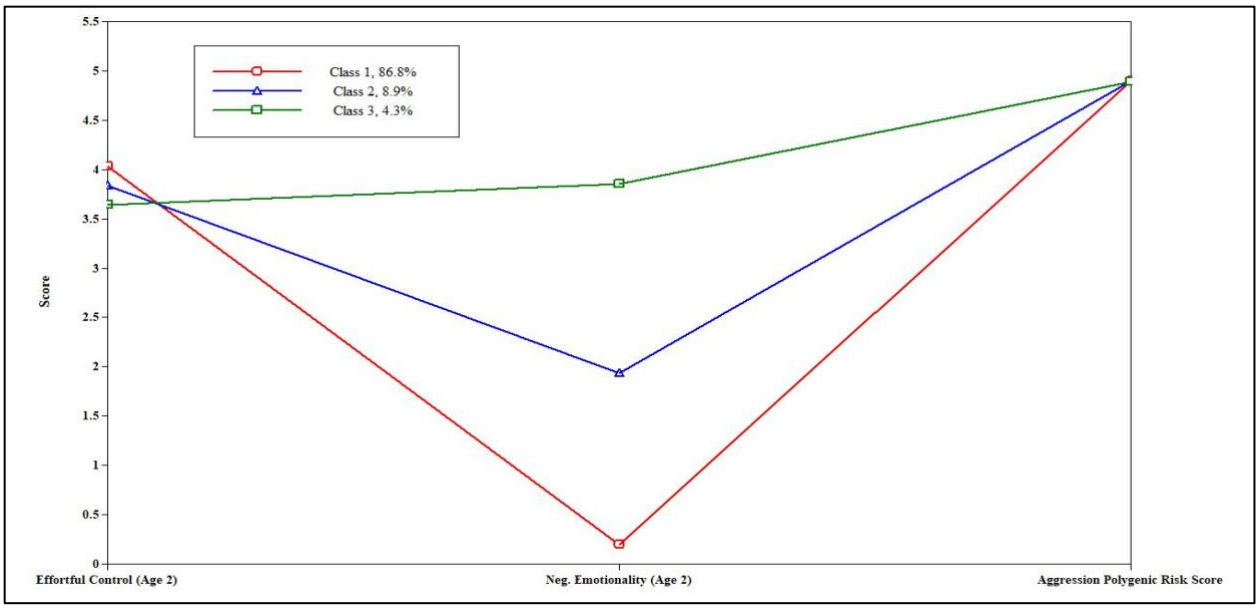

Figure S1. Three profile solution: Latent profiles of observed negative emotionality at age 2, parent-reported effortful control at age 2, and aggression genome-wide polygenic scores.

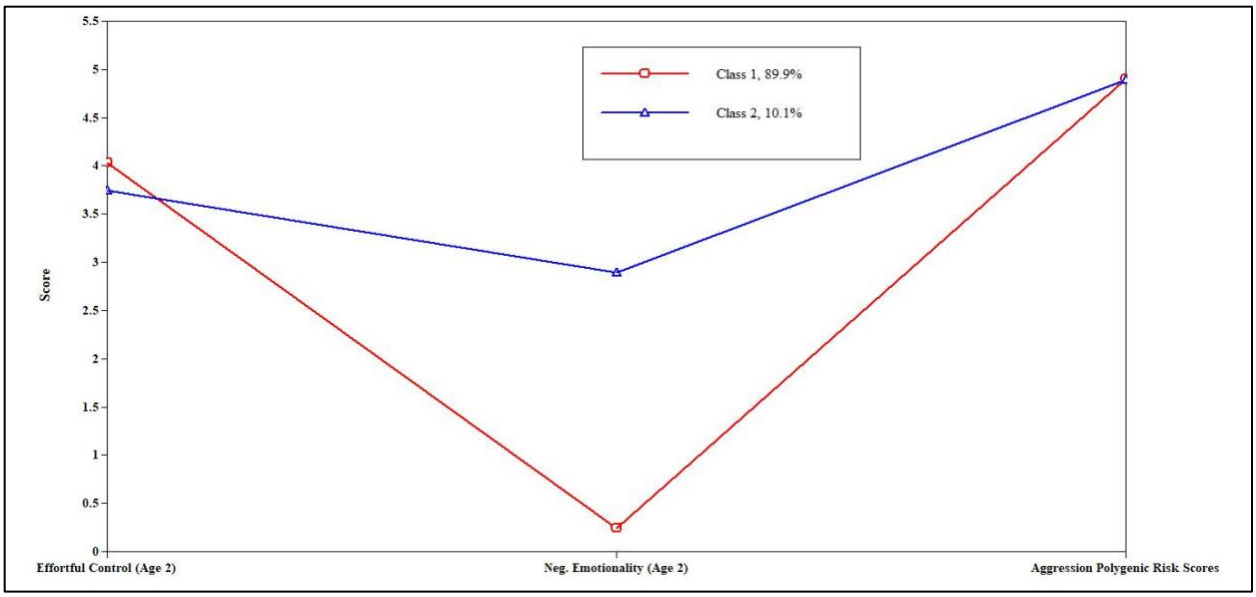

Figure S2. Two profile solution: Latent profiles of observed negative emotionality at age 2, parent-reported effortful control at age 2, and aggression genome-wide polygenic scores.

BIBLIOGRAPHY

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

Ahadi, S. A., Rothbart, M. K., & Ye, R. (1993). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality, 7*(5), 359–378. https://doi.org/10.1002/per.2410070506

Aron, E. N., & Aron, A. (1997). Sensory-processing sensitivity and its relation to introversion and emotionality. *Journal of Personality and Social Psychology*, *73*(2), 345–368. http://dx.doi.org/10.1037/0022-3514.73.2.345

Aron, E. N., Aron, A., & Jagiellowicz, J. (2012). Sensory processing sensitivity: A review in the light of the evolution of biological responsivity. *Personality and Social Psychology Review, 16*(3), 262-282. https://doi.org/10.1177/1088868311434213

Andruff, H., Carraro, N., Thompson, A., & Gaudreau, P. (2009). Latent class growth modelling: A tutorial. *Tutorials in Quantitative Methods for Psychology*, *5*(1), 11–24. https://doi.org/10.20982/tqmp.05.1.p011

Auerbach, J. G., Benjamin, J., Faroy, M., Geller, V., & Ebstein, R. (2001). DRD4 related to infant attention and information processing: A developmental link to ADHD? *Psychiatric Genetics*, *11*(1), 31–35.

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2006). Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. Developmental Psychobiology, 48(5), 406-409. https://doi.org/10.1002/dev.20152

Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2011). Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. *Development and Psychopathology, 23*(1), 39-52. https://doi.org/10.1017/S0954579410000635

Beauchaine, T. P., Gartner, J., & Hagen, B. (2000). Comorbid depression and heart rate variability as predictors of aggressive and hyperactive symptom responsiveness during inpatient treatment of conduct-disordered, ADHD boys. *Aggressive Behavior, 26*(6), 425-441. https://doi.org/10.1002/1098-2337(200011)26:6<425:AID-AB2>3.0.CO;2-I

Beaver, J. D., Lawrence, A. D., Van Ditzhuijzen, J., Davis, M. H., Woods, A., & Calder, A. J. (2006). Individual differences in reward drive predict neural responses to images of food.

The Journal of Neuroscience, 26(19), 5160-5166. https://doi.org/10.1523/JNEUROSCI.0350-06.2006

Belsky, J. (1997a). Variation in susceptibility to environmental influence: An evolutionary argument, *Psychological Inquiry*, *8*(3), 182–186. https://doi.org/10.1207/s15327965pli0803_3

Belsky, J. (1997b). Theory testing, effect-size evaluation, and differential susceptibility to rearing influence: The case of mothering and attachment. *Child Development, 68*, 598–600. https://doi.org/10.2307/1132110

Belsky, J. (2005). Differential susceptibility to rearing influence: An evolutionary hypothesis and some evidence. In B. Ellis & D. Bjorklund (Eds.), *Origins of the social mind: Evolutionary Psychology and Child Development*, (pp. 139-163). NY: Guildford.

Belsky, J., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. Current Directions in Psychological Science, 16(6), 300–304. https://doi.org/10.1111/j.1467-8721.2007.00525.x

Belsky, D. W., & Israel, S. (2014). Integrating Genetics and Social Science: Genetic Risk Scores. Biodemography and Social Biology, 60(2), 137–155. https://doi.org/10.1080/19485565.2014.946591

Belsky, J., Hsieh, K.-H., & Crnic, K. (1996). Infant positive and negative emotionality: One dimension or two? *Developmental Psychology*, *32*(2), 289–298. https://doi.org/10.1037/0012-1649.32.2.289

Belsky, J., Hsieh, K.-H., & Crnic, K. (1998). Mothering, fathering, and infant negativity as antecedents of boys' externalizing problems and inhibition at age 3 years: Differential susceptibility to rearing experience? *Development and Psychopathology, 10*(2), 301-319. http://dx.doi.org/10.1017/S095457949800162X

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, *135*(6), 885–908. https://doi.org/10.1037/a0017376

Belsky, J., & Pluess, M. (2013a). Beyond risk, resilience and dysregulation: phenotypic plasticity and human development. *Development and Psychopathology, 25*, 1243-1261. https://doi.org/10.1017/S095457941300059X

Belsky, J., & Pluess, M. (2013b). Genetic moderation of early child-care effects on social functioning across childhood: A developmental analysis. *Child Development, 84*(4), 1209–1225. https://doi.org/10.1111/cdev.12058

Belsky, J., & Pluess, M. (2012). Differential susceptibility to long-term effects of quality of child care on externalizing behavior in adolescence? *International Journal of Behavioral Development*, *36*(1), 2–10. https://doi.org/10.1177/0165025411406855

Bradley, R., & Corwyn, R. (2008). Infant temperament, parenting, and externalizing behavior in first grade: A test of the differential susceptibility hypothesis. *Journal of Child Psychology and Psychiatry, 49*(2), 124-131. https://doi.org/10.1111/j.1469-7610.2007.01829.x

Bradley, R., Corwyn, R. F., McAdoo, H. P., & García Coll, C. (2001). The home environments of children in the United States Part I: Variations by age, ethnicity, and poverty status. *Child Development, 72*(6), 1844-1867. https://doi.org/10.1111/1467-8624.t01-1-00382

Burt, S. A., & Mikolajewski, A. J. (2008). Preliminary evidence that specific candidate genes are associated with adolescent-onset antisocial behavior. *Aggressive Behavior, 34*(4), 437-445. https://doi.org/10.1002/ab.20251

Carey, W. B. (1973). Measuring infant temperament in pediatric practice. In J. C. Westrnan (Ed.), *Individual differences in children*. New York, NY: Wiley.

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification, 13*(2), 195–212. https://doi.org/10.1007/BF01246098

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Crnic, K. A., & Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development, 61*(5), 1628-1637. http://dx.doi.org/10.2307/1130770

Connell, A., Bullock, B. M., Dishion, T. J., Shaw, D., Wilson, M., & Gardner, F. (2008). Family intervention effects on co-occurring early childhood behavioral and emotional problems: A latent transition analysis approach. *Journal of Abnormal Child Psychology, 36*(8), 1211-1225. https://doi.org/10.1007/s10802-008-9244-6

De La Vega, F. M., & Bustamante, C. D. (2018). Polygenic risk scores: A biased prediction? *Genome Medicine, 10*(1), 100. https://doi.org/10.1186/s13073-018-0610-x

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The Validity of the Multi-Informant Approach to Assessing Child and Adolescent Mental Health. *Psychological Bulletin, 141*(4), 858–900. https://doi.org/10.1037/a0038498

Dishion, T. J., Brennan, L. M., Shaw, D. S., McEachern, A. D., Wilson, M. N., & Jo. B. (2014). Prevention of problem behavior through annual Family Check-Ups in early childhood: Intervention effects from the home to early elementary school. *Journal of Abnormal Child Psychology, 42*(3), 343–354. https://doi.org/10.1007/s10802-013-9768-2

Dishion, T. J., Gardner, K., Patterson, G. R., Reid, J. B., & Thibodeaux, S. (1983). The Family Process Code: A multidimensional system for observing family interaction. *Unpublished Coding Manual. Oregon Social Learning Center;*, (January 1983), 92426-97401. Unpublished coding manual.

Dishion, T. J., Hogansen, J., Winter, C., & Jabson, J. M. (2004). *The Coder Impressions Inventory*. Unpublished coding manual. Available from the Child and Family Center, 6217 University of Oregon, Eugene, OR 97403.

Dishion, T. J., Shaw, D. S., Connell, A., Wilson, M. N., Gardner, F., & Weaver, C. (2008). The Family Check Up with high-risk families with toddlers: Outcomes on positive parenting and early problem behavior. *Child Development, 79*, 1395-1414.

Dishion, T. J., & Stormshak, E. A. (2007). *Intervening in children's lives: An ecological, family-centered approach to mental health care*. Washington, DC, US: American Psychological Association. http://dx.doi.org/10.1037/11485-000

Dishion, T. J., Stormshak, E. A., & Kavanagh, K. A. (2012). Everyday parenting: A professional's guide to building family management skills. Research Press.

Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Losoya, S., Murphy, B., … Reiser, M. (2000). Prediction of Elementary School Children's Externalizing Problem Behaviors from Attentional and Behavioral Regulation and Negative Emotionality. *Child Development, 71*(5), 1367–1382. https://doi.org/10.1111/1467-8624.00233

Eisenberg, N., Spinrad, T. L., Fabes, R. A., Reiser, M., Cumberland, A., Shepard, S. A., … Murphy, B. (2004). The Relations of Effortful Control and Impulsivity to Children's Resiliency and Adjustment. *Child Development, 75*(1), 25–46.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining delinquency and drug use*. Thousand Oaks, CA: Sage.

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2011). Differential susceptibility to the environment: An evolutionary--neurodevelopmental theory. *Development and Psychopathology, 23*(1), 7–28. https://doi.org/10.1017/S0954579410000611

Euesden, J., Lewis, C. M., O'Reilly P. F., PRSice: Polygenic Risk Score software, *Bioinformatics 31*(9), 1466–1468, https://doi.org/10.1093/bioinformatics/btu848

Everitt, B. J., & Robbins, T. W. (2000). Second-order schedules of drug reinforcement in rats and monkeys: Measurement of reinforcing efficacy and drug-seeking behaviour. *Psychopharmacology,153*(1), 17-30. https://doi.org/10.1007/s002130000566

Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2004). Parent-Adolescent Disagreement Regarding Psychopathology in Adolescents From the General Population as a Risk Factor for Adverse Outcome. *Journal of Abnormal Psychology, 113*(2), 198–206. https://doi.org/10.1037/0021-843X.113.2.198

Forbes, E. E., Brown, S. M., Kimak, M., Ferrell, R. E., Manuck, S. B., & Hariri, A. R. (2009). Genetic variation in components of dopamine neurotransmission impacts ventral striatal reactivity associated with impulsivity. *Molecular Psychiatry, 14*(1), 60–70.

Ganzel, B. L., & Morris, P. A. (2011). Allostasis and the developing human brain: Explicit consideration of implicit models. *Development and Psychopathology, 23*(4), 955–974. https://doi.org/10.1017/S0954579411000447

Gardner, F. (2000). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review, 3*(3), 185-198. https://dx.doi.org/1 0.1023/A:1009503409699

Goldsmith, H. H., Buss, A. H., Plomin, R., Rothbart, M. K., Thomas, A., Chess, S., . . . McCall, R. B. (1987). Roundtable: What is temperament? Four approaches. *Child Development, 58*, 505–529. https://doi.org/10.2307/1130527

Granic, I., Hollenstein, T., Dishion, T. J., & Patterson, G. R. (2003). Longitudinal Analysis of Flexibility and Reorganization in Early Adolescence: A Dynamic Systems Study of Family Interactions. https://doi.org/10.1037/0012-1649.39.3.606

Guo, G., Roettger, M. E., & Shih, J. C. (2007). Contributions of the DAT1 and DRD2 genes to serious and violent delinquency among adolescents and young adults. *Human Genetics, 121*(1), 125–136. https://doi.org/10.1007/s00439-006-0244-8

Halpern, L. F., Garcia Coll, C. T., Meyer, E. C., & Bendersky, K. (2001). The contributions of temperament and maternal responsiveness to the mental development of small-for gestational-age and appropriate-for-gestational-age infants. *Journal of Applied Developmental Psychology, 22*(2), 199-224. http://dx.doi.org/10.1016/S01933973(01)00077-6

Hariri, A.R., Brown, S.M., Williamson, D.E., Flory, J.D., de Wit, H., & Manuck, S.B. (2006). Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *The Journal of Neuroscience, 26*(51), 13213-13217.

Hollenstein, T. (2007). State space grids: Analyzing dynamics across development. *International Journal of Behavioral Development, 31*(4), 384–396. https://doi.org/10.1177/0165025407077765

Hollenstein, T., Granic, I., Stoolmiller, M., & Snyder, J. (2004). Rigidity in parent-child interactions and the development of externalizing and internalizing behavior in early childhood. *Journal of Abnormal Child Psychology*. https://doi.org/10.1023/B:JACP.0000047209.37650.41

Holmboe, K., Nemoda, Z., Fearon, R. M. P., Sasvari-Szekely, M., & Johnson, M. H. (2011). Dopamine D4 receptor and serotonin transporter gene effects on the longitudinal development of infant temperament. *Genes, Brain, and Behavior, 10*(5), 513–522. https://doi.org/10.1111/j.1601-183X.2010.00669.x

Ivorra, J. L., D'Souza, U. M., Jover, M., Arranz, M. J., Williams, B. P., Henry, S. E., … Molto, M. D. (2011). Association between neonatal temperament, SLC6A4, DRD4 and a functional polymorphism located in TFAP2B. *Genes, Brain and Behavior, 10*(5), 570–578. https://doi.org/10.1111/j.1601-183X.2011.00696.x

Jabson, J., Dishion, T. J., Gardner, F., & Burton, J. (2004). Relationship Process Code v-2.0 Training Manual: A system for coding relationship interactions. In *Unpublished coding manual.* Eugene, Oregon: Child and Family Center, University of Oregon;

Jung, T., & Wickrama, K. A. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Kahn, R. S., Khoury J., Nichols, W. C., & Lanphear, B. P. (2003). Role of dopamine transporter genotype and maternal prenatal smoking in childhood hyperactive, impulsive, inattentive, and oppositional behaviors. *The Journal of Pediatrics, 143*(1), 104-110. https://doi.org/10.1016/S0022-3476(03)00208-7

Karg, K., Burmeister, M., Shedden, K., & Sen, S. (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: Evidence of genetic moderation. *Archives of General Psychiatry, 68,* 444–454. http://dx.doi.org/10.1001/archgenpsychiatry .2010.189

Karreman , A., De Haas, S., Van Tuijl, C., Van Aken, M. A., & Deković, M. (2010). Relations among temperament, parenting and problem behavior in young children. *Infant Behavior and Development, 33*(1), 39-49. https://doi.org/10.1016/j.infbeh.2009.10.008

Keller, M. C. (2014). Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) Solution. *Biological Psychiatry*, *75*. https://doi.org/10.1016/j.biopsych.2013.09.006

Kendler, K. S. (2006). "A gene for...": the nature of gene action in psychiatric disorders. *Focus,162*(3), 1243-400.

Keers, R., Coleman, J. R. I., Lester, K. J., Roberts, S., Breen, G., Thastum, M., … Eley, T. C. (2016). A Genome-Wide Test of the Differential Susceptibility Hypothesis Reveals a Genetic Predictor of Differential Response to Psychological Treatments for Child Anxiety Disorders. *Psychotherapy and Psychosomatics*, *85*(3), 146–158. https://doi.org/10.1159/000444023

Kiff, C. J., Lengua, L. J., & Bush, N. R. (2011). Temperament variation in sensitivity to parenting: Predicting changes in depression and anxiety. *Journal of Abnormal Child Psychology, 39*(8), 1199-1212. https://doi.org/10.1007/s10802-011-9539-x

Kim, M. S., Patel, K. P., Teng, A. K., Berens, A. J., & Lachance, J. (2018). Genetic disease risks can be misestimated across global populations. *Genome Biology*, *19*(1), 179. https://doi.org/10.1186/s13059-018-1561-7

Kochanska, G., Aksan, N., & Joy, M. E. (2007). Children's fearfulness as a moderator of parenting in early socialization: Two longitudinal studies. *Developmental Psychology, 43*(1), 222-237. http://dx.doi.org/10.1037/0012-1649.43.1.222

Kochanska, G., & Kim, S. (2013). Early attachment organization with both parents and future behavior problems: From infancy to middle childhood. *Child Development, 84*(1), 283-296. http://doi.org/10.1111/j.1467-8624.2012.01852.x

Lamey, A., Hollenstein, T., Lewis, M. D., & Granic, I. (2004). GridWare (Version 1.1). [Computer software]. http://statespacegrids.org/

Lemery-Chalfant, K., Clifford, S., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (in press). Genetic moderation of the effects of the Family Check-Up intervention on children's internalizing symptoms: A longitudinal study with a racial/ethnically diverse sample. *Development and Psychopathology 30*(5). https://doi.org/10.1017/S095457941800127X

Lengua, L. J. (2008). Anxiousness, frustration, and effortful control as moderators of the relation between parenting and adjustment in middle-childhood. *Social Development, 17*(3), 554-577. https://doi.org/10.1111/j.1467-9507.2007.00438.x

Lunkenheimer, E. S., Dishion, T. J., Shaw, D. S., Connell, A. M., Gardner, F., Wilson, M. N., & Skuban, E. M. (2008). Collateral benefits of the family check-up on early childhood school readiness: Indirect effects of parents' positive behavior support. *Developmental Psychology*, *44*(6), 1737–1752. https://doi.org/10.1037/a0013858

Manuck, S.B. (2011a). *Delay discounting covaries with childhood socioeconomic status as a function of genetic variation in the dopamine D4 receptor (DRD4).* Paper presented at the Society for Research in Child Development, Montreal, Canada.

Manuck, S. B., & McCaffery, J. M. (2014). Gene-environment interaction. *Annual Review of Psychology, 65*(1), 41-70. https://doi.org/10.1146/annurev-psych-010213-115100

McEachern, A. D., Fosco, G. M., Dishion, T. J., Shaw, D. S., Wilson, M. N., & Gardner, F. (2013). Collateral benefits of the family check-up in early childhood: Primary caregivers' social support and relationship satisfaction. *Journal of Family Psychology, 27*(2), 271.

Meunier, J. C., Roskam, I., & Browne, D. T. (2011). Relations between parenting and child behavior: Exploring the child's personality and parental self-efficacy as third variables. *International Journal of Behavioral Development, 35*(3), 246-259. https://doi.org/10.1177/0165025410382950

Miller, W., & Rollnick, S. (2003). Motivational interviewing: Preparing people for change. *Journal for Healthcare Quality*, *25*(3), 46.

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology*, *13*(2), 355–372. https://doi.org/10.1017/S0954579401002097

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. In *Development and Psychopathology* (Vol. 14).

Moilanen, K. L., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. (2009). Predictors of Longitudinal Growth in Inhibitory Control in Early Childhood. *Soc Dev*, *19*(2), 326–347. https://doi.org/10.1111/j.1467-9507.2009.00536.x.Predictors

Mooney, M. A., & Wilmot, B. (2015). Gene Set Analysis: A Step-By-Step Guide. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, *168*(7), 517–527. https://doi.org/10.1002/ajmg.b.32328

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychological Bulletin*, *110*(3), 406–425. https://doi.org/10.1037/0033-2909.110.3.406

Moran, L. R., Lengua, L. J., & Zalewski, M. (2013). The Interaction between Negative Emotionality and Effortful Control in Early Social-emotional Development: Negative Emotionality and Effortful Control. *Social Development*, *22*(2), 340–362. https://doi.org/10.1111/sode.12025

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. https://doi.org/10.1111/j.1530-0277.2000.tb02070.x

Muthén, L. K., & Muthén, B. (2012). 1998-2012. *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Narvaes, R., & Martins de Almeida, R. M. (2014). Aggressive behavior and three neurotransmitters: dopamine, GABA, and serotonin—A review of the last 10 years. *Psychology & Neuroscience, 7*(4), 601–607. http://dx.doi.org/10.3922/j.psns.2014.4.20

Nikolaidis, A., & Gray, J., R. (2010). ADHD and the DRD4 exon III 7-repeat polymorphism: An international meta-analysis. *Social Cognitive and Affective Neuroscience, 5*(2-3), 188-193. https://doi.org/10.1093/scan/nsp049

Nikolova, Y. S., Ferrell, R. E., Manuck, S. B., & Hariri, A. R. (2011). Multilocus genetic profile for dopamine signaling predicts ventral striatum reactivity. *Neuropsychopharmacology, 36*(9), 1940–1947. http://doi.org/10.1038/npp.2011.82

Noldus Information Technology (2012). The Observer XT Reference Manual (Version 11.0). [Computer software]. https://www.noldus.com/human-behavior-research/products/the-observerxt?gclid=EAIaIQobChMIpN2qy_zc4wIVFZSzCh1z6w0DEAAYASAAEgJMP_D_BwE

Olson, S. L., Sameroff, A. J., Kerr, D. C. R., Lopez, N. L., & Wellman, H. M. (2005). Developmental foundations of externalizing problems in young children: The role of effortful control. *Development and Psychopathology*, *17*(01), 25–45. https://doi.org/10.1017/S0954579405050029

Oniszczenko, W., & Dragan, W. L. (2005). Association Between Dopamine D4 Receptor Exon III Polymorphism and Emotional Reactivity as a Temperamental Trait. *Twin Research and Human Genetics*, *8*(6), 633–637. https://doi.org/10.1375/twin.8.6.633

Pappa, I., St Pourcain, B., Benke, K., Cavadino, A., Hakulinen, C., Nivard, M. G., ... & Evans, D. M. (2016). A genome-wide approach to children's aggressive behavior: The EAGLE consortium. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *171*(5), 562-572.

Peterson, J., Winter, C., Jabson, J., & Dishion, T. J. (2008). Relationship Affect Coding System. Unpublished manuscript, Child and Family Center, University of Oregon, Eugene, OR.

Pitzer, M., Jennen-Steinmetz, C., Esser, G., Schmidt, M. H., & Laucht, M. (2011). Differential susceptibility to environmental influences: The role of early temperament and parenting in the development of externalizing problems. *Comprehensive Psychiatry, 52*(6), 650-658. https://doi.org/10.1016/j.comppsych.2010.10.017

Pluess, M., & Belsky, J. (2010). Differential susceptibility to parenting and quality child care. *Developmental Psychology, 46*(2), 379-390. https://doi.org/10.1037/a0015203

Pluess, M., & Belsky, J. (2013). Vantage sensitivity: Individual differences in response to positive experiences. *Psychological Bulletin, 139*(4), 901-916. http://dx.doi.org/10.1037/a0030196

Poehlmann, J., Hane, A., Burnson, C., Maleck, S., Hamburger, E., & Shah, P. E. (2012). Preterm infants who are prone to distress: Differential effects of parenting on 36-month behavioral and cognitive outcomes. *Journal of Child Psychology and Psychiatry, 53*(10), 1018-1025. https://doi.org/10.1111/j.1469-7610.2012.02564.x

Poehlmann, J., Schwichtenberg, A. J., Shlafer, R. J., Hahn, E., Bianchi, J. P., & Warner, R. (2011). Emerging self-regulation in toddlers born preterm or low birth weight: Differential susceptibility to parenting? *Development and Psychopathology, 23*(1), 177-193. https://doi.org/10.1017/S0954579410000726

Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature, 538*(7624), 161–164. https://doi.org/10.1038/538161a

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*(4), 437–448. https://doi.org/10.3102/10769986031004437

Propper, C., Willoughby, M., Halpern, C. T., Carbone, M. A., & Cox, M. (2007). Parenting quality, DRD4, and the prediction of externalizing and internalizing behaviors in early childhood. *Developmental Psychobiology, 49*(6), 619-632. https://doi.org/10.1002/dev.20249

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., A. R. Ferreira, M., Bender, D., … C Sham, P. (2007). PLINK: A tool set for whole-genome association and population-based linkage

analyses. *American Journal of Human Genetics, 81*, 559–575. https://doi.org/10.1086/519795

Qadeer, M. I., Amar, A., Mann, J. J., Hasnain S. (2017) Polymorphisms in dopaminergic system genes; association with criminal behavior and self-reported aggression in violent prison inmates from Pakistan. *PLoS ONE, 12*(6), E0173571. https://doi. org/10.1371/journal.pone.0173571

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Reid, J. B. (1993). Prevention of conduct disorder before and after school entry: Relating interventions to developmental findings. *Development and Psychopathology, 5*, 243-262. http://dx.doi.org/10.1017/s0954579400004375

Reid, P., Monsen, J. J., & Rivers, I. (2004). Psychology's contribution to understanding and managing bullying within schools. *Educational Psychology in Practice, 20*(3), 241-258. https://doi.org/10.1080/0266736042000251817

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), 490-529. http://dx.doi.org/10.1037/0033-2909.128.3.490

Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science, 273*(5281), 1516-1517. https://doi.org/10.1126/science.273.5281.1516

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology, 9*(1), 22-29. http://dx.doi.org/10.1080/15374418009532938

Roisman, G. I., Newman, D. A., Fraley, R. C., Haltigan, J. D., Groh, A. M., & Haydon, K. C. (2012). Distinguishing differential susceptibility from diathesis–stress: Recommendations for evaluating interaction effects. *Development and Psychopathology, 24*(2), 389-409. https://doi.org/10.1017/S0954579412000065

Rolstad, S., Adler, J., & Rydén, A. (2011). Response burden and questionnaire length: Is shorter better? A review and meta-analysis. *Value in Health, 14*(8), 1101–1108. https://doi.org/10.1016/j.jval.2011.06.003

Rothbart, M. K., & Ahadi, S. A. (1994). Temperament and the development of personality. Journal of Abnormal Psychology, 103(1), 55-66. http://dx.doi.org/10.1037/0021-843X.103.1.55

Rothbart, M., A. Ahadi, S., L. Hershey, K., & Fisher, P. (2001). Investigations of Temperament at Three to Seven Years: The Children's Behavior Questionnaire. *Child Development, 72*, 1394–1408. https://doi.org/10.1111/1467-8624.00355

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon & N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 105-176). Hoboken, NJ, US: John Wiley & Sons Inc.

Rothbart, M. K., & Bates, J. E. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 99-166). Hoboken, NJ, US: John Wiley & Sons Inc.

Reid, J. B. (1993). Prevention of conduct disorder before and after school entry: Relating interventions to developmental findings. *Development and Psychopathology, 5*, 243-262. http://dx.doi.org/10.1017/s0954579400004375

Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: Multiple varieties but real effects. *Journal of child Psychology and Psychiatry, 47*(3-4), 226-261. https://doi.org/10.1111/j.1469-7610.2005.01557.x

Sameroff, A. J. (1983). Resilient Children and How They Grew. *PsycCRITIQUES, 28*(1). https://doi.org/10.1037/021523

Sameroff, A. J., Seifer, R., & Elias, P. K. (1982). Sociocultural variability in infant temperament ratings. *Child Development*, 164-173.

Shaw, D. S., Connell, A., Dishion, T. J., Wilson, M. N., & Gardner, F. (2009). Improvements in maternal depression as a mediator of intervention effects on early childhood problem behavior. *Development and Psychopathology*, *21*(2), 417-439. https://doi.org/10.1017/S0954579409000236

Shaw, D.S., Dishion, T. J., Supplee, L. H., Gardner, F., & Arnds, K. (2006). A family-centered approach to the prevention of early-onset antisocial behavior: Two-year effects of the family check-up in early childhood. *Journal of Consulting and Clinical Psychology*, *74*, 1-9.

Shaw, D. S., Galán, C., Lemery-Chalfant, K., Dishion, T. J., & Wilson, M. N. (2019). Early predictors of children's early-starting conduct problems: Child, family, genetic, and intervention effects. *Development and Psychopathology.*

Shaw, D. S., Sitnick, S. L., Brennan, L. M., Choe, D. E., Dishion, T. J., Wilson, M. N., &

Gardner, F. (2016). The long-term effectiveness of the Family Check-Up on school-age conduct problems: Moderation by neighborhood deprivation. *Development and Psychopathology*, *28*, 1471–1486. https://doi.org/10.1017/S0954579415001212

Shaw, D. S., Supplee, L., Dishion, T. J., Gardner, F., & Arnds, K. (2006). Randomized trial of a family-centered approach to the prevention of early conduct problems: 2-Year effects of the family check-up in early childhood. *Journal of Consulting and Clinical Psychology, 74*(1),1. https://doi.org/10.1037/0022-006X.74.1.1

Sitnick, S. L., Shaw, D. S., Weaver, C. M., Shelleby, E. C., Choe, D. E., Reuben, J. D., … Taraban, L. (2017). Early Childhood Predictors of Severe Youth Violence in Low-Income Male Adolescents. *Child Development*, *88*(1), 27–40. https://doi.org/10.1111/cdev.12680

Slagt, M., Dubas, J. S., Deković, M., & van Aken, M. A. (2016). Differences in sensitivity to parenting depending on child temperament: A meta-analysis. *Psychological Bulletin*, *142*(10), 1068.

Slagt, M., Dubas, J. S., Ellis, B. J., Aken, M. A. G. van, & Deković, M. (2019). Linking emotional reactivity "for better and for worse" to differential susceptibility to parenting among kindergartners. *Development and Psychopathology*, *31*(2), 741–758. https://doi.org/10.1017/S0954579418000445

Sweitzer, M. M., Halder, I., Flory, J. D., Craig, A. E., Gianaros, P. J., Ferrell, R. E., & Manuck, S. B. (2013). Polymorphic variation in the dopamine D4 receptor predicts delay discounting as a function of childhood socioeconomic status: Evidence for differential susceptibility. *Social Cognitive and Affective Neuroscience*, *8*(5), 499–508. https://doi.org/10.1093/scan/nss020

Uher, R., & McGuffin, P. (2010). The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. *Molecular Psychiatry, 15*, 18–22. https://dx.doi.org/10.1038/mp.2009.123

van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2015). Genetic differential susceptibility on trial: Meta-analytic support from randomized controlled experiments. *Development and Psychopathology, 27*(1), 151-162.

Vaughn, B. E., Taraldson, B. J., Crichton, L., & Egeland, B. (1981). The assessment of infant temperament: A critique of the Carey Infant Temperament Questionnaire. *Infant Behavior & Development, 4*(1), 1-17. http://dx.doi.org/10.1016/S0163-6383(81)80003-3

Vermunt, J. K., Vermunt, & K., J. (2010). Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. *Political Analysis*, *18*(04), 450–469.

Weeland, J., Slagt, M., Brummelman, E., Matthys, W., De Castro, B. O., & Overbeek, G.(2015). 5-HTTLPR expression outside the skin: An experimental test of the emotional reactivity hypothesis in children. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0141474

Vitaro, F., Barker, E. D., Boivin, M., Brendgen, M., & Tremblay, R. E. (2006). Do early difficult temperament and harsh parenting differentially predict reactive and proactive aggression? *Journal of Abnormal Child Psychology, 34*(5), 681-691. https://doi.org/10.1007/s10802-006-9055-6

Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. A., Dudbridge, F., & Middeldorp, C. M. (2014). Research review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry, 55*(10), 1068-1087. https://doi.org/10.1111/jcpp.12295

Zohsel, K., Buchmann, A. F., Blomeyer, D., Hohm, E., Schmidt, M. H., Esser, G., . . . Laucht, M. (2014). Mothers' prenatal stress and their children's antisocial outcomes--a moderating role for the dopamine D4 receptor (DRD4) gene. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *55*(1), 69-76. https://doi.org/10.1111/jcpp.12138