Genetic, cognitive and intervention effects on parenting, child attachment and children's psychobiology

Runze, Jana

2024

DOI (link to publisher)
10.5463/thesis.543

document version
Publisher's PDF, also known as Version of record

Genetic, cognitive and intervention effects on parenting, child attachment and children's psychobiology

Jana Runze

Cover illustration and design by Sonja Lugtenberg
Layout design by Sonja Lugtenberg
Printed by Ridderprint, www.ridderprint.nl

Please cite as: Runze, J. (2024). Genetic, cognitive and intervention effects on parenting, child attachment and children's psychobiology. (Doctoral Dissertation). Amsterdam: Department of Clinical Child and Family Studies, Vrije Universiteit Amsterdam, The Netherlands.

VRIJE UNIVERSITEIT

GENETIC, COGNITIVE AND INTERVENTION EFFECTS ON PARENTING, CHILD ATTACHMENT AND CHILDREN'S PSYCHOBIOLOGY

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op vrijdag 22 maart 2024 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

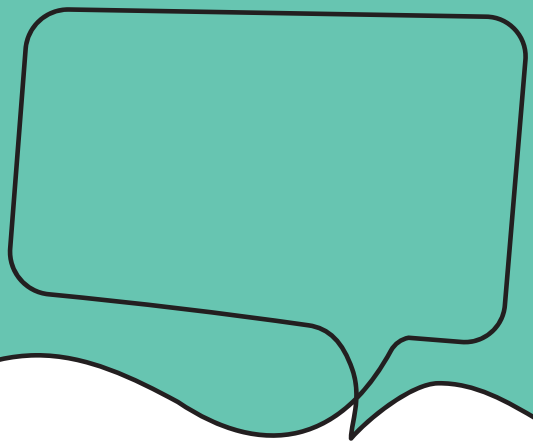
Jana Runze

geboren te Gummersbach, Duitsland

promotor:	prof.dr. M.J. Bakermans-Kranenburg
copromotor:	dr. M. Oosterman
promotiecommissie:	prof.dr. M. Dekovic prof.dr. M. Dozier prof.dr. R.M.P. Fearon prof.dr. G.J. Overbeek prof.dr. P.S. Sterkenburg

Table of Contents

Chapter 1	General Introduction	7
Chapter 2	Response Bias is Genetically Biased: Another argument for Kagan's Philippic against Questionnaires in Developmental Psychology	25
Chapter 3	The Polygenic and Reactive Nature of Observed Parenting	47
Chapter 4	Actigraphic sleep and cortisol in middle childhood: A multivariate behavioral genetics model	71
Chapter 5	Replicating a randomized trial with video-feedback to promote positive parenting in parents of school-aged twins	99
Chapter 6	Conduct Problems and Hair Cortisol Concentrations Decrease in School-Aged Children after VIPP-SD: A Randomized Controlled Trial in Two Twin Cohorts	125
Chapter 7	Differential Susceptibility in the Intergenerational Transmission of Secure Base Script Knowledge?	149
Chapter 8	General Discussion	169
Chapter 9	References	186
	Author contribution Statements	224
	Summary	226
	List of Publications	229
	Curriculum Vitae	231
	PhD Portfolio	232
	Acknowledgements	234



GENERAL INTRODUCTION

CHAPTER ONE

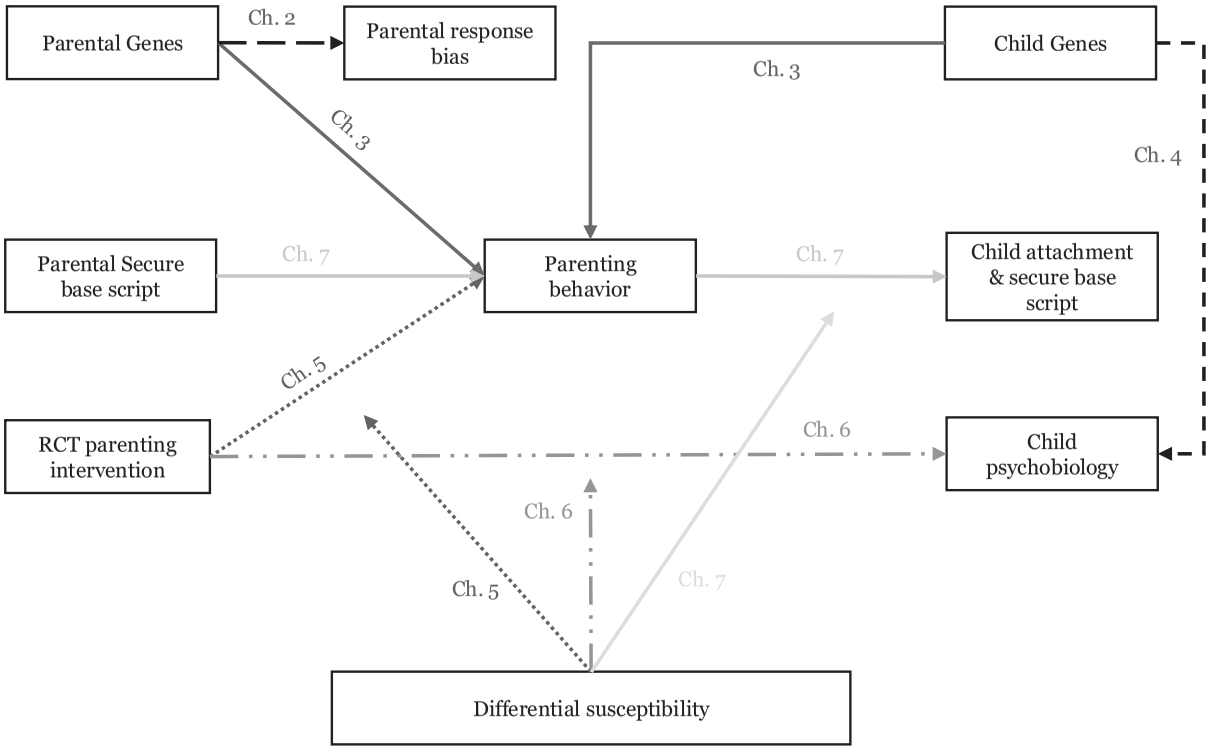
1

Parents, and their multifaceted relationship with and behavior towards their children play a major role in the development of a child (Belsky, 1984). The development of children does not only hold significance for the children themselves, but also for society since children grow up to be active members of our society. Problems in the psychobiological development of children have been linked to (mental) health difficulties later in life and subsequent personal, social and economic costs (Bor et al., 2004). Preventing the onset of (mental) health difficulties by supporting adaptive child development is therefore highly desired.

In this thesis, I will first examine whether a genetic component underlies responding in a biased way to questionnaire items. Then, I will shed light on parents as a driving factor in the psychobiological development of children and investigate the genetic and environmental determinants of parents' behavior towards and relationship with their offspring. My dissertation revolves around the process model of parenting (developed by Belsky in 1984, adapted by Taraban and Shaw in 2018) as well as the intergenerational transmission model of attachment (Verhage et al., 2016). Figure 1.1 shows the determinants of parenting behavior that are the focus of this thesis: parents' genes, their children's genes, parents' secure base scripts, and a randomized controlled trial (RCT) with a parenting intervention. Moreover, I will shed light on determinants of children's attachment and their psychobiology, focusing on children's genes, their parents' behavior and the effect of the parenting intervention.

Figure 1.1

Visualization of the thesis' main topics: determinants of parenting and parenting effects on the child



Reliability of self- or parent-report questionnaires

Traditionally, questionnaires have served as valuable tools in psychological research. Questionnaires are an easy, cost-efficient way of gathering data, yet their reliability has raised significant concerns, particularly in light of the limited correspondence observed between self-reported responses and empirical observations (Prince et al., 2008). Over thirty years ago, Van Dam and Van IJzendoorn (1988) revealed a lack of overlap between parents' self-reports of their infants' attachment security and the assessments conducted by trained raters employing the Strange Situation Procedure. Ambiguity in questionnaire items may be one of the reasons for this disparity. Even well-established instruments, such as the widely employed 'Strengths and Difficulties Questionnaire' (SDQ; Goodman, 1997), have been found to possess invalidity due to the presence of several ambiguous items. For instance, consider this item: “I am nervous in new situations. I easily lose confidence.” In a study of Vrijhof et al. (2016), this question was split in two sets of disambiguated items. These items in turn exhibited a substantial mean difference of half a standard deviation, with a mere 20% overlapping variance (Vrijhof et al., 2016).

Moreover, response biases might play a role. The inclination of individuals to respond to questions with inherent biases poses a significant challenge. For example, the acquiescence bias is a common response tendency of participants wherein they tend to agree to questions regardless of their actual behavior or emotions, possibly due to a desire to conform (Ray, 1983). This bias can distort questionnaire results by creating systematic error. It is not entirely clear why people differ in their inclination to respond in a biased way. One possibility is that response biases might be partially explained by genetic factors, although in-depth examination of this hypothesis is lacking. If genetic predispositions indeed play a role in response biases, this underscores the necessity of augmenting traditional questionnaire-based research with more objective and biologically grounded methodologies, including observations, controlled experiments, and the incorporation of genetic or neurobiological measures.

Genetic factors in parenting

To explore the role of genetics in parenting several approaches have been

utilized in the past. The use of the classical twin design has been a practical option. The foundation for the classical twin design lies in the notion that monozygotic twins (MZ) are genetically identical, i.e., they share 100% of their genes, whereas dizygotic twins (DZ) are like full siblings, i.e. they share, on average, 50% of their genes. The classical twin design allows us to decompose the total phenotypic variance into variance explained by additive genes (A), variance explained by the common, shared environment (C) and unique environmental factors and measurement error (E), commonly called the "ACE" model (Neale et al., 2016). The environmental variance is divided into "shared" and "unique", because we assume that twins who are raised together share a certain amount of their environment, such as the neighborhood they grow up in. However, they may be subject to distinctive environmental factors as well, such as one twin participating in football and the other twin in gymnastics (Neale et al., 2016). There has been limited research into parenting using twin designs because twin samples are often composed of twins-as-children and not twins-as-parents (Klahr & Burt, 2014). In a meta-analysis of six parent-based behavioral genetic studies, the authors found that the unique environment accounted for 63–90% of the variance in parenting, depending on the parenting dimension (Klahr & Burt, 2014) and probably on how the phenotype was measured (observed versus self-report measures).

The candidate gene approach has been used to provide insight into the genetic mechanisms underlying differences in parenting. In a hypothesis driven method, possible candidate genes of parenting have been detected and examined which centered around three neurotransmitter systems: dopamine, oxytocin and serotonin. For example, the 10-repeat polymorphism of the dopamine transporter DAT1 has been found to be associated with more negative maternal behaviors (Lee et al., 2010). Moreover, the rs53576 polymorphism on the oxytocin receptor gene (OXTR) and the short allele on the serotonin transporter gene 5HTT were found to be associated with parental sensitive responsiveness (Bakermans-Kranenburg & Van IJzendoorn, 2008). However, these early studies yielded conflicting results, and replicating the main effects of candidate genes appeared challenging (Bakermans-Kranenburg & Van IJzendoorn, 2014; Mileva-Seitz et al., 2016). It has been argued that effects of early candidate gene studies might have been overestimated due to

the winner's curse phenomenon (Lohmueller et al., 2003). According to the winner's curse phenomenon which is based on auction theory, the winning bid is likely to overestimate the value of an item and a significant finding in candidate gene studies can be seen as a winning bid. Genetic effect size estimates are based on the significant variants found, which results in upwardly biased effect size estimates leading to replication studies being underpowered and failing to corroborate the significant findings.

A promising innovation in the field of genetics of parenting research has been the introduction of polygenic scores (PGSs) based on genome-wide association studies (GWAS). GWAS investigate the entire genome in relatively large samples to identify associations between single-nucleotide polymorphisms (SNPs) and a certain phenotype. The results can be used to summarize the effects of millions of genetic variants across the genome into a single polygenic score which captures the genetic predisposition to a particular phenotype (Dudbridge, 2013). PGSs hold vast potential as a novel approach as they reflect the polygenicity of complex phenotypes and enable the quantification of a person's individual genetic proclivities (Plomin & von Stumm, 2018).

However, because polygenic scores are based on GWAS, their quality and usefulness depend on the quality of the GWAS and the phenotypic measure used in it. One of the auspicious polygenic scores is the PGS of educational attainment (Lee et al., 2018) as the PGS is based on a GWAS with a large sample size and the PGS shows acceptable predictive power of educational achievements (i.e., the PGS of educational attainment is able to predict actual educational attainment). Moreover, research suggests that these scores are associated with parental cognitive and behavioral skills which are in turn associated with parenting (Crandall et al., 2015). Indeed, in the Dunedin Study with observed parenting and parental GWAS data in 702 participants (Wertz et al., 2019), parents with a higher PGS for educational attainment (PGS-EA, Lee et al., 2018) provided more sensitive parenting to their 3-year-old child. Additionally, Wertz and colleagues (2020) replicated these findings in the E-Risk Longitudinal twin study with 1116 families, wherein mothers with higher PGS-EA scores provided more sensitive parenting towards their 5-12 year old children. Besides the PGS-EA, the PGS of IQ (Savage et al., 2018) and the PGS

of income (Hill et al., 2019) might be plausible candidates for parenting research, as they are based on a GWAS with large sample size and show acceptable predictive power of their respective phenotypes. Moreover, the phenotypes of intelligence, income and educational attainment have been found to be correlated but the PGSs have not been explored in the context of parenting (Ritchie & Tucker-Drob, 2018; Yang & Qiu, 2016; Zagorsky, 2007). Thus, while there are some first preliminary findings regarding a potential role of polygenic scores in parenting research, the studies are in need of replication and extension.

Gene-environment correlations

Although investigating genetic and environmental effects separately is necessary, the interaction of both should also be considered, as behavioral genetics research indicates that they are correlated or interact with each other (Avinun & Knafo, 2014; Klahr & Burt, 2014). Gene-environment correlations can be classified into passive, evocative and active correlations (Plomin et al., 1977). Passive gene-environment correlations originate in parents' genes. For example, parental genes may impact the degree to which the child's environment will be cognitively stimulating, consequently influencing cognitive development beyond the genetic predisposition inherited by the child (Wertz et al., 2019). Evocative gene-environment effects indicate that greater genetic similarity in children is related to more similar received parenting. In families with MZ versus DZ twins more similar parenting behavior was found, (Euser et al., 2020), and a meta-analysis provided evidence for substantial evocative gene-environment correlations (Avinun & Knafo, 2014). We speak of active gene-environment correlation when individuals choose certain environments based on their genetic composition. Apart from gene-environment correlations, gene-environment interactions can also occur. A special case of gene-environment interaction is the differential susceptibility hypothesis which will be discussed in a later paragraph.

Secure base scripts in parenting

One of the potential factors why parents differ in their parenting abilities besides their genetic make-up might be their own attachment experiences in the

past. Parents' experiences with their own attachment figures are organized into a cognitive schema: the secure base script (Waters & Waters, 2006). Parents who were subjected to consistent and sensitive caregiving develop complete, consolidated and easily accessible secure base scripts. However, individuals who have been exposed to inconsistent or insensitive caregiving develop incomplete, inconsistent or ineffective knowledge of the secure base script (Nivison et al., 2021; Waters & Waters, 2006). These secure base scripts guide parents' behavior in their interactions with their offspring, such as sensitive caregiving and sensitive discipline (Waters & Roisman, 2019; Witte et al., 2023) and have also been found to be associated with child attachment (Verhage et al., 2016).

Child attachment

Bowlby (1982) was the first to develop attachment theory stating that all children establish an attachment relationship with their parents and depending on the type of attachment relationship exhibit various attachment behaviors to elicit support and comfort. Securely attached children use their caregiver as a secure base from which they feel comfortable exploring their environment (Ainsworth et al., 1978). Insecurely attached children, in contrast, have difficulties relying on their caregiver in times of need. Children with an insecure attachment are classified into one of three types: avoidant, ambivalent, and disorganized (Ainsworth et al., 1978; Main & Hesse, 1990). Children with an avoidant attachment are avoiding to turn to their caregiver for support in times of need whereas children with an ambivalent attachment have difficulties being comforted by the support of their caregivers and often show passivity or anger. Finally, disorganized children display conflicted or apprehensive behavior toward their caregiver in situations of need (Main & Hesse, 1990).

The type of attachment relationship might be influenced by the quality of care the child receives from their caregiver over multiple interactions (De Wolff & Van IJzendoorn, 1997). Whereas securely attached children tend to have sensitive parents, insecurely attached children tend to have harsh, inconsistent or rejecting parents (De Wolff & Van IJzendoorn, 1997; Verhage et al., 2016). Being sensitive as a parent has been defined as being able to notice and interpret the child's signals

accurately and to respond to these signals appropriately and promptly (Ainsworth et al., 1978). In contrast, inconsistent or rejecting parenting is defined as being inconsistent in one's parenting behaviors and showing rejecting behaviors or engaging in harsh disciplining behaviors, such as verbal aggression (e.g. yelling and frequent negative commands) and physical aggression (e.g. hitting) (Chang et al., 2003; Rohner et al., 2008).

Intergenerational transmission of attachment

As stated before, parental attachment representations have been linked to their offspring's attachment relationship. Previous research has suggested that this transmission of attachment is mediated by parental sensitivity (Verhage et al., 2016). The association between parental attachment representations and child-parent attachment via parental sensitivity is known as intergenerational transmission of attachment (Verhage et al., 2016). However, there is a transmission gap between parental attachment representations and child attachment, i.e., the concordance between parents' and children's attachment cannot be fully explained by observed parental behavior. More recent and unpublished studies reported weaker effect sizes for transmission. One explanation for the transmission gap is the (unexplored) role of moderating factors, such as differential susceptibility of either the parents or the children, which I will come back to in a later paragraph.

Interventions to improve parenting

According to the process model of parenting (Belsky, 1984; Taraban & Shaw, 2018), how parents behave depends not only on their genetic make-up, but also on psychosocial factors. A multitude of psychosocial factors have been extensively studied and found to be associated with parenting, such as depression (Lovejoy et al., 2000), social support (Green et al., 2007), attachment representations (Verhage et al., 2016) or marital quality (Krishnakumar & Buehler, 2000). Consequently, parental behavior is amenable to interventions. A well-established parenting support program is the Video-feedback Intervention to promote Positive Parenting (VIPP-SD), which draws from social learning and attachment theory to enhance parental sensitivity and sensitive discipline. Meta-analytical evidence has shown that the VIPP-SD is

effective in increasing parental sensitivity and parental sensitive discipline (Juffer et al., 2017; van IJzendoorn et al., 2022). During the VIPP-SD, parents learn to respond adequately to difficult child behavior, such as with distraction and inductive discipline. Furthermore, parents learn to recognize subtle signals of their child and to react swiftly and appropriately. The intervention is standardized and individualized, and has been evaluated in diverse groups (Van IJzendoorn et al., 2022). In a recent meta-analysis of 25 randomized controlled VIPP-SD studies, the effects of the intervention on parental sensitivity, sensitive discipline and secure child-parent attachment have been corroborated (Van IJzendoorn et al., 2022). VIPP-SD has originally been developed for young children aged 1-3 and its effectiveness has been studied in pre-school children as well, but effects in families with older children still had to be examined.

Genetic factors in child psychobiology

Knowledge about the nature of psychobiological factors and whether these are predominantly determined by genes or the environment is essential for informing parenting interventions. Understanding the extent to which sleep patterns and cortisol levels are genetically determined or shaped by the environment provides the possibility to personalize intervention approaches which consider a child's genetic predispositions. Sleep is one of the important factors to consider because it has been found to have an impact on the psychological and physiological functioning of children (Hatzinger et al., 2013; Sadeh et al., 2014). Two physiological mechanisms are involved in sleep regulation: the suprachiasmatic nucleus and the sleep homeostat. The suprachiasmatic nucleus is the biological clock indicating that it is time to sleep when it is dark and time to be awake when it is light (Moore, 2007), while the sleep homeostat keeps track of how much sleep we had and when we must refill our sleep reserve (Borbély & Achermann, 1999). A recent twin study on sleep revealed that genetic factors accounted for 81% of the variance in sleep duration and for 79% of the variance in sleep efficiency in a group of 8-year-old children (Breitenstein et al., 2021). A different picture is provided by twin studies on cortisol. Cortisol is a glucocorticoid regulated by the hypothalamus-pituitary-adrenal axis (HPA-axis). Cortisol is involved in several processes, such as the immune system,

the regulation of the metabolism, and the response to stress (Oakley & Cidlowski, 2013). In stressful conditions, higher amount of cortisol is released to cope with the stressor (E. Russell et al., 2012). In a study on 14-year-olds, heritability of the cortisol level after morning awakening was 28% (Ouellet-Morin et al., 2016). It is uncertain, whether differences in sleep and cortisol can be explained by the same genetic or environmental factors. Both, sleep and cortisol regulation, are intricately connected within the physiological systems of the human body. For example, genetic variations may influence the production of neurotransmitters such as serotonin or dopamine which are involved in both sleep and cortisol regulation (Nakamaru-Ogiso et al., 2012; G. Russell & Lightman, 2019). Understanding these shared genetic factors can provide valuable insights into how changes in one system (e.g. cortisol levels) might have spillover effects on the other system (e.g. sleep). Interventions focusing on improving environmental factors such as parenting have the potential to play an important role in the stress system and/or sleep regulation.

Parenting's impact on child psychobiology

Unsurprisingly, aside from potential genetic factors, parenting has been found to predict children's broader socio-emotional and cognitive development. Both parental sensitivity and child-parent attachment have been found to be related to prosocial behavior (Deneault et al., 2023; Leerkes et al., 2009), emotional intelligence (Walker et al., 2022) and executive functioning (Borairi et al., 2021). Lastly, meta-analyses have shown that attachment and parental sensitivity are protective factors in the occurrence of internalizing and externalizing behavior problems (Cooke et al., 2022; Deneault et al., 2021; Groh et al., 2017).

Although parenting was successfully improved using the VIPP-SD, in the meta-analysis previously discussed, Van IJzendoorn and colleagues (2022) did not find significant evidence for effects of the parenting intervention on children's externalizing behavior. The inconsistency of results warrants further investigation to provide robust evidence for the effectiveness of attachment-based parenting intervention on children's behavior. Furthermore, parenting has been found to influence children's psychobiology, particularly their cortisol levels.

Previous parenting interventions based on attachment theory were successful

in decreasing diurnal or hair cortisol levels in children (Bakermans-Kranenburg et al., 2008; Poehlmann-Tynan et al., 2020), indicating that parenting also gets "under the skin" by affecting cortisol secretion of the HPA-axis. However, in a meta-analysis of 18 parenting intervention studies, no overall significant effect of parenting interventions on child cortisol levels was found (Martins 2020). However, most of the studies did not report the effectiveness of the parenting intervention on parenting behaviors. Moreover, most samples were small and methodological procedures were evaluated as poor. Only one study investigated hair cortisol levels which are more indicative of chronic stress as compared to salivary cortisol level which provide a measure of acute cortisol levels and are influenced by many different confounding factors. Further work is needed to provide more robust evidence for the effectiveness of attachment-based parenting interventions on child externalizing behavior and stress regulation and to investigate mechanisms revolving around parents as a driving factor in the psychobiological development of children.

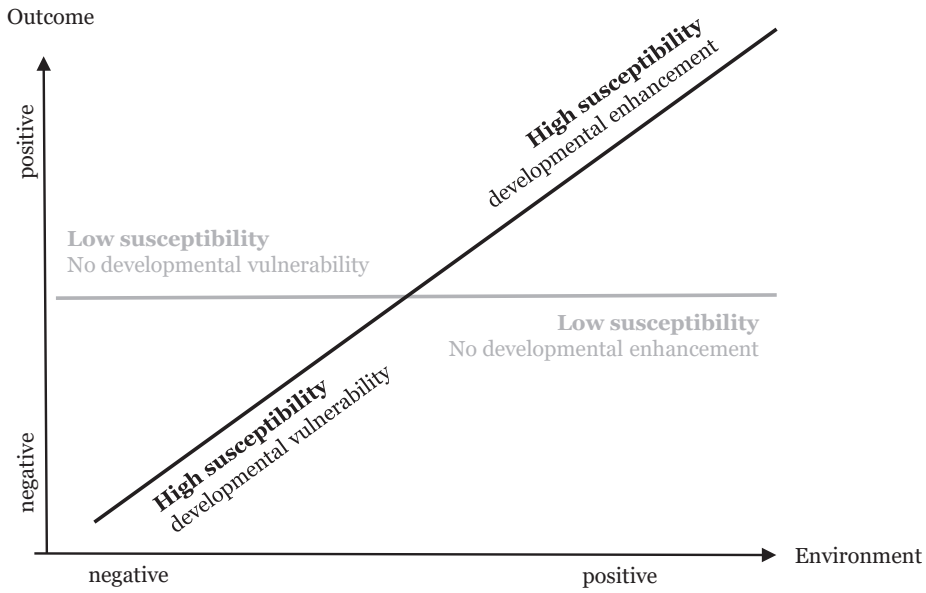
Differential susceptibility

The inconclusive results of parenting intervention effects on child development as well as the transmission gap of attachment might partly be explained by the differential susceptibility hypothesis. The differential susceptibility hypothesis arises from the concept that some individuals might be more (or less) affected by environmental factors (Belsky et al., 2007). In a similar theory, the diathesis-stress theory, this assumption holds true only for negative influences, i.e., that susceptible individuals "suffer" more from negative experiences (Ellis et al., 2011). In the differential susceptibility, the susceptibility is assumed to hold for negative and positive experiences. For example, when receiving a parenting intervention, parents who are more susceptible to the environment might profit more from the intervention compared to parents who are less susceptible (Bakermans-Kranenburg & Van IJzendoorn, 2015). Likewise, highly susceptible children may, when exposed to negative parenting, show more psychological problems, but might thrive more in the presence of positive parenting (Belsky et al., 2007; Ellis et al., 2011). Being affected by a certain environment is advantageous only when the future environment is predictable. However, this is rarely true as the future can be highly uncertain (e.g.

COVID crisis with isolation, invasion of Ukraine by Russia, but also in terms of quality of relationships, employment or welfare). Therefore, it is not known which kind of family environment (for example which parenting strategy) will provide the best fit to ensure the child's adaptive development in life. Belsky (2005) presented the concept of "bet-hedging" against an uncertain future: For families it would be beneficial to have children that are differentially susceptible to the environment, with some that would thrive in a future environment that is similar to the early environment and others that are less affected by influences such as parental behavior and efforts to shape child behavior. Figure 1.2 shows the moderating role of susceptibility.

Figure 1.2

Association between environment and outcome in the differential susceptibility theory



Note. Adapted from Bakermans-Kranenburg and Van IJzendoorn (2015)

Various indicators of differential susceptibility have been used in previous research, commonly dimensions of temperament, such as fearfulness, negative reactivity, impulsivity, anger proneness, withdrawal, negativity or sensitivity (Slagt et al., 2016). Some genetic variables (such as the DRD4 7-repeat allele, the 5-HTTLPR

allele and, more recently, polygenic scores of differential susceptibility) have also been explored as susceptibility markers (Bakermans-Kranenburg & Van IJzendoorn, 2015; Belsky & Van IJzendoorn, 2017; Keers et al., 2016).

For example, in several studies, children with negative reactivity have been found to show more internalizing and externalizing behavior as well as worse social and academic adjustment compared to children with low negative reactivity when parenting quality was low, but they showed less problematic behavior and fewer adjustment difficulties when parenting quality was high (Ellis et al., 2011). Differential susceptibility assessed by polygenetic variables has been explored by Keers and colleagues (2016) who were the first to develop a GWAS-based PGS for differential susceptibility to anxiogenic environments. Based on within-pair variability in emotional problems in monozygotic twins, they identified SNPs of environmental susceptibility and combined these in a PGS. This PGS was found to significantly moderate the effects of parenting on child emotional problems (Keers et al., 2016).

However, the findings regarding susceptibility using different measures of exposures, outcomes and differential susceptibility factors have been inconclusive as some studies did not find any effects (Euser et al., 2021; Slagt et al., 2016). For example, parents scoring high on temperamental reactivity did not profit more from a parenting intervention compared to parents scoring low on temperamental reactivity. Furthermore, their parenting qualities were also not worse in the control condition compared to the parents with low temperamental reactivity (Euser et al., 2021). One explanation might be that differential susceptibility may be domain-specific (Boyce & Ellis, 2005; Ellis et al., 2011). For example, Zhang et al. (2021) found that individual differences in susceptibility to family-social effects were not associated with individual differences in susceptibility to quality-of-care cognition-related effects: Individuals who were highly susceptible to the family environment with regard to social outcomes were not susceptible in the domain of quality-of-care with regard to cognitive outcomes. Moreover, a recent paper by Belsky et al. (2022) showed that around half of the children who were highly susceptible to childcare quality effects on pre-academic skills were not highly susceptible to childcare quantity effects on behavior problems and vice versa. This suggests that differential susceptibility is

domain-specific instead of domain-generic: Indicators of differential susceptibility might only be reliable indicators when examining certain environments or specific outcomes. All in all, research examining domain-specificity of susceptibility is still in its early stages. More research is warranted to establish domain-specific, reliable polygenic markers of differential susceptibility and disentangle which marker reflects differential susceptibility in which environmental contexts.

Aims

My dissertation has four aims: The first aim is to examine the potential role of genes in the response bias in questionnaire research (Chapter 2). The second aim of my dissertation is to investigate the determinants of parenting (Chapter 3, Chapter 5, Chapter 6 and Chapter 7). The third aim of my dissertation is to examine the role of genes and parenting in child attachment and psychobiology (Chapter 4, Chapter 6 and Chapter 7). Finally, the fourth aim is to examine the role of differential susceptibility in the context of parent-child relationships (Chapter 5, Chapter 6 and Chapter 7).

Setting

This dissertation is based on two studies. Most studies were performed using data from the Leiden Consortium on Individual Development (L-CID, www.developmentmatters.nl). L-CID is a longitudinal cohort-sequential study including parents and their twin children. Data was collected from families in two cohorts: the early childhood cohort with 238 families and the middle childhood cohort with 257 families. Six yearly visits took place wherein a variety of measures were collected, including observational measures of parenting, biological measures such as saliva or hair samples and questionnaires (for more information see Crone et al., 2020). In addition, I made use of data from the Generation R (GenR) study which is a population-based prospective cohort study based in Rotterdam. In total, 9778 pregnant mothers were enrolled in order to investigate early environmental and genetic correlates of normal and abnormal development (for more information see Kooijman et al., 2016).

Outline

In **Chapter 2**, I examined the potential role of genes in response bias of parents and their twin children. In **Chapter 3**, I investigated whether genetic predispositions to socio-economic and cognitive factors (i.e. educational attainment, IQ and income) can predict sensitive parenting in mothers of the GenR study. In **Chapter 4,** I investigated the heritability of sleep and cortisol, in order to examine the potential for environmental (i.e., parenting) effects on sleep and cortisol of their children. In **Chapter 5**, I moved on to psychosocial determinants of parenting. I examined whether a randomized controlled trial in the form of a brief attachment-based video-feedback intervention (VIPP-SD) increased parental sensitivity and sensitive discipline in parents of school-aged twin children. **Chapter 6** expands the research about VIPP-SD by investigating its effect on children's hair cortisol levels and conduct problems. In **Chapter 7**, I examined whether the transmission of attachment representation is mediated by parental sensitivity and sensitive discipline and whether children are differentially susceptible to parental sensitivity and sensitive discipline. In **Chapter 8**, findings of the previous chapters are discussed.

This chapter is published as:

Runze, J. & Van IJzendoorn, M.H. (2023). Response Bias is Genetically Biased: Another argument for Kagan's Philippic against Questionnaires in Developmental Psychology. *Developmental Psychology*. doi: 10.1037/dev0001614. Epub ahead of print. PMID: 37902678.

RESPONSE BIAS IS GENETICALLY BIASED: ANOTHER ARGUMENT FOR KAGAN'S PHILIPPIC AGAINST QUESTIONNAIRES IN DEVELOPMENTAL PSYCHOLOGY

CHAPTER TWO

Abstract

One of the *Trio of Concerns* of Jerome Kagan (2007) was the frequent use of questionnaires in developmental psychology and related disciplines. His main reasons were the minimal overlap between (self-) reported and observed phenotypes, the ambiguity of items, and systematic socio-economic status disparities in responding. We wondered whether genetic differences would also influence response bias in case of ambiguous or even absurd items triggering an acquiescence bias to agree with whatever impossible content of the question. We examined the genetic predisposition for the acquiescence response bias in 257 families with twins. Both parents and twins completed a modified Wildman Symptoms Questionnaire and provided salivary DNA samples. From published Genome-Wide Association Studies (GWAS) we derived polygenic score (PGS) algorithms for indicators of socio-economic status (educational attainment and income) and for IQ, which were applied to the GWAS results of our participants' DNA. The twins (N = 514 children, 55% monozygotic pairs) were on average 7.92 years old (SD = 0.66), 51.6% were female. The primary parents were on average 40.48 years old (SD = 4.66), 91% were female. In 90% of the families, both parents were born in the Netherlands and 91% were from a middle or higher socio-economic background. Higher polygenic scores for educational attainment and income but not for IQ predicted less acquiescence response bias in parents and children. In addition to Kagan's reasons for his concerns about questionnaires we found a genetic predisposition to response bias triggered by the ambiguous form of questions and answers.

One of the concerns Jerome Kagan (2007) discussed in his thought-provoking essay *A Trio of Concerns* was the widely spread use of questionnaires in developmental psychology and related disciplines. He complained about the uncritical acceptance of popular personality self-reports such as the Big Five (NEO, Costa & McCrae, 1985) and questionnaires for parents about their children's behavior problems such as the CBCL (Achenbach & Rescorla, 2000) because they were used as if they delivered valid descriptions of real personalities or actual behaviors. Respondents however want to present themselves or their children in a favorable light, questions rarely are unambiguous, and several studies show low correlations between different self-reports for the same characteristics, as well as low correspondence between observed and self-reported behaviors.

In a caustic critique of one of our own studies on the early childhood aggression curve (Alink et al., 2006) he dismissed our rationalization to use the CBCL instead of behavioral observations. We used the outworn argument of limited (time and budgetary) resources which he curtly rejected with the following analogy: "A chemist who asked three expert informants to estimate the chemical composition of a white powder, rather than perform the time-consuming analysis necessary to determine the ingredients, would be a target of satire or serious criticism." (Kagan, 2007, p.370).

Did Kagan have valid reasons to express his serious concerns about the excessive use of questionnaires in developmental science? His main technical reasons were the minimal overlap between (self-) reported and observed data on developmental phenotypes, and the ambiguity of many items in widely used questionnaires, and systematic socio-economic status disparities in responding. Related to the first issue, in one of our earliest studies on attachment we found no overlap whatsoever between parents' reports of their infants' attachment security on the one hand and attachment security observed by trained raters in the Strange Situation Procedure on the other hand (van Dam & van IJzendoorn, 1988).

Regarding the second issue, we found that even well-established questionnaires, such as the famous and still widely used 'Strengths and Difficulties Questionnaire' (SDQ; Goodman, 1997) which targets externalizing and internalizing behavior problems and prosociality, and is cited more than 10,000 times, are invalid because of several ambiguous items. By splitting ambiguous items such as "I am

nervous in new situations. I easily lose confidence" in two unambiguous questions we found that the two sets of disambiguated items showed a mean difference of half a standard deviation, with only 20% overlapping variance (Vrijhof et al., 2016).

Forty years ago, one of the authors asked 175 first year university students to complete a 'questionnaire' without questions but with response options such as: 'very satisfied, satisfied, unsatisfied, very unsatisfied' and 'always, sometimes, never' (van IJzendoorn, 1984). Remarkably, almost all students completed the questionnaire (or better: 'optionaire') but even more surprising was the finding that the distributions of the chosen responses deviated significantly from chance, with large effect sizes in favor of 'satisfied' or 'sometimes'. The answers to the questionnaire without questions were not completely arbitrary but shaped by formal features of the response options, with a preference for going middle of the road or for avoiding very negative responses. The acquiescence response bias creates systematic errors and is a potential confounder that might inflate associations between independent and dependent variables. Most worrisome would be a genetic component of this response bias suggesting that some respondents are genetically more inclined than others to let themselves be directed by the bias.

In the current study we developed an acquiescence response bias questionnaire with impossible or even absurd items to examine the genetics of endorsing such items. We modified the Wildman Symptoms Questionnaire (Wildman & Wildman, 1999) for use in developmental research and examined the heritability of the acquiescence response bias in the answers of twin children and their parents to impossible developmental questions, such as: "My child grows so fast that I can sometimes see them grow" (see Supplement 1 for all items of the Wildman Developmental Questionnaire). Using the behavior genetics approach we explored the difference between monozygotic and dizygotic twins in within-twin pair correlation of the acquiescence response bias, computing the additive genetic estimate and the contribution of the shared and unique environmental influences to the response bias (the ACE model; Bouchard Jr & Propping, 1993).

In the molecular genetics approach we used polygenic scores (PGSs) derived from recent large-scale genome-wide association studies (GWAS) to determine the molecular genetic component of the bias in the children and their parents.

Polygenic scores are increasingly used in developmental research to assess an individual's genetic predisposition for a wide range of complex physical and mental traits and diseases. A polygenic score aggregates the (regression) effects of multiple genetic variants, or single nucleotide polymorphisms (SNPs), found in GWAS to be associated with a particular trait or disease. These genetic variants may each have tiny individual effects, but combined in a polygenic score, they might explain a substantial part of the variance in the phenotypes. A great advantage of PGSs is the unequivocal temporal order and thus potential causal role they play in explaining individual phenotypic differences as genetic variants are pre-dating exposure to environmental influences in the individual life course. PGSs may not be causal of differences in human development in any direct sense however because many pathways through gene expression, proteins and neuronal connections come in-between but they can nevertheless shape individual differences (Plomin, 2019).

Because Kagan (2007) mentioned social class as an important correlate of responses to questionnaires with ambiguous and unintelligible content we selected PGSs for social class (educational attainment and income) and added a PGS of intelligence to estimate a causal role of these genetic scores in the acquiescence response bias. A large genetic component explaining variance in the Wildman Developmental Questionnaire responses might be interpreted as evidence for a bias to agree with content in incomprehensible items. From published GWAS studies we derived the weights for polygenic scores of educational attainment (Lee et al., 2018), IQ (Savage et al., 2018), and family income (Hill et al., 2019), and we computed the polygenic scores for each of our study participants, parents and children. We also tested whether phenotypic school attainment, IQ and income are associated with the corresponding PGSs and the Wildman scores, and whether the phenotypes predicted variance in the Wildman scores above and beyond the PGSs.

Methods

Design

The L-CID project is a longitudinal study including parents with their twin children (www.developmentmatters.nl). Families with twins from the western region of the Netherlands received an invitation letter and an information brochure.

Parents who indicated a willingness to participate received a phone call to check the inclusion criteria and to provide more information about the study. Families were eligible if the twins had the same sex, their parents were able to communicate in Dutch and parents as well as grandparents were born in Europe. The occurrence of congenital disability, psychological disorder, chronic illness, hereditary disease, visual/hearing impairment, or an IQ of <70 led to exclusion. Six yearly visits were planned and the current study utilized data from the third, fourth, fifth and sixth assessment (T3, T4, T5, T6; collected from 2018-2022). Approval for this study was provided by the central committee on research involving human subjects (CCMO; Middle childhood cohort NL50277.058.14).

Participants

Through an invitation letter, 1174 families were contacted of which N = 257 families with twins were enrolled in the study (N = 514 children, 55% monozygotic pairs). At enrollment, the children were on average 7.92 years old (SD = 0.66) and 51.6% were female. The primary parents were on average 40.48 years old (SD = 4.66) and 91% were female. In 90% of the families, both parents were born in the Netherlands and 91% of the families were from a middle or high socio-economic background (based on parental education). For the current study, utilizing data from waves T3 to T6, 248 parents and 460 children provided data.

Measures and Procedure

IQ. Children completed the third edition of the Wechsler Intelligence Scale for Children (WISC-III, Wechsler, 1991) at T3. The WISC-III is an individually administered intelligence test with three composite IQ scores, the full-scale IQ, a verbal IQ and a performance IQ. The test is suited for children between 6 and 16 years of age.

Wildman Developmental Questionnaire. The Wildman Symptoms Questionnaire (Merckelbach et al., 2009; Wildman & Wildman, 1999) was modified by focusing the questions on fake or highly implausible developmental phenomena. The Wildman Developmental Questionnaire (from now on called: Wildman) consists of 5 items scored on a 5-point Likert-scale (ranging from 1 = not at all to

5 = extremely). An example item is: "When I am stressed, I can feel it in my feet". Children answered these questions for themselves whereas parents answered them for their children (e.g. "When my child is stressed, they can feel it in their feet", see Supplement 1 for all items). Parents and their children filled in the questions at all time points (T3, T4, T5, T6) and reliability was acceptable (α = .70 for the children and α = .88 for the parents).

Polygenic scores. Saliva samples were collected from the primary parent and the children at T2 and subsequently genotyped and imputed at the Genetic Laboratory of the Department of Internal Medicine (Population Genomics) at Erasmus MC. A complete description of the process can be found in Supplement 1. We computed three polygenic scores: (1) educational attainment (EA), (2) IQ, and (3) income (see Table S2.1 for information about the base samples). For all PGSs we used the PRSice software (Choi & O'Reilly, 2019). The GWAS summary statistics served as the base sample, and L-CID was the target sample. Only autosomal SNPs were used, since there is no consensus for the sex chromosomes (Choi & O'Reilly, 2019). PGSs were calculated using clump r^2=0.1, 250kb at different p-value thresholds (i.e., 1, .50, .30, .20, .10, .05, and .001). The optimal p-value threshold is not known a priori and depends on several factors, such as power of the base and target data and the effect size distribution. Preselection of an arbitrary p-value threshold would lead to underperformance of the PGS prediction. Therefore, we performed regression analyses with all PGSs under the different p-value thresholds as predictors. Subsequently, the PGSs under the best p-value threshold were selected and subsequently used for further analyses (Choi et al., 2020) (see Table S2.2).

Data analysis

For all analyses, we used the statistical software R (version 4.2.2, R Core Team, 2022). We assessed missingness using the mcar test of the naniar package (Tierney et al., 2022). Missingness was not completely at random as indicated by Little's MCAR test (for parents: X^2 (400) = 554.55, p < .001; for children: X^2 (340) = 830.01, p < .001). One of the reasons for high missingness is that DNA samples were collected at T2 whereas the informed consent for the actual analysis of the data was asked in T5, when several participants had dropped out. Participants who dropped

out at T5 did not differ from participants who were included in this study with regards to age, sex, psychiatric diagnoses or IQ. Data were imputed using multiple imputation as recommended in the case of non MCAR data (Newman, 2014). We imputed missing data using the package mice (Van Buuren & Groothuis-Oudshoorn, 2011) with 50 iterations for each of the 5 imputed datasets using predictive mean matching for all variables in our statistical models. Following the guidelines of Wulff and Ejlskov (2017), we also report results based on 47 imputed datasets with 20 iterations in the supplementary materials (Table S2.12). We included age of the child at all measurement time points, gender of the parent, ethnicity of the child, socio-economic status, IQ of the child, handedness of the child and zygosity of the child as auxiliary variables. Even in cases of large proportions of missingness, MI reduces bias and is preferable above other techniques of handling missing data (Madley-Dowd et al., 2019). However, we also report sensitivity analyses with complete cases.

First, we conducted univariate and bivariate genetic covariance structure models using the umx package (Bates et al., 2019) to estimate the contributions of genetic and environmental factors to the phenotypic variance of our variables of interest. The genetic covariance structure model included the contributions of additive genetic factors (A), shared environmental factors (C), and unique environmental factors including measurement error (E). To estimate these contributions, we exploited the fact that MZ twins share 100% of their alleles (i.e., are genetically identical), while DZ twin on average share 50% of their alleles. Consequently, MZ twins will show greater resemblance than DZ twin with respect to phenotypes that are subject to genetic influences. We first fitted the saturated model, in which we estimated the (unconstrained) MZ and DZ covariance matrices. This saturated model served as a baseline model. Subsequently, we fitted the ACE model. We conducted the statistical tests using the log-Likelihood Ratio Test (LRT) statistic and adopted an alpha of .05.

Second, we conducted hierarchical structural equation models to investigate the association between the PGSs and the responses on the Wildman in parents using the lavaan package (Rosseel, 2012). To investigate the responses of the twin children on the Wildman, we were not able to use multilevel modelling, as there was no variance in most of the PGS variables within the twin pairs (due to shared

genetics). Therefore we randomly divided the twin pairs up into two separate datasets and conducted hierarchical structural equation models separately for each twin. We report model fit, which was assessed with the Comparative Fit Index (CFI; Bentler, 1990), the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973), and the Root Mean Square Error of Approximation (RMSEA; Hu & Bentler, 1998). However, indices can be below/above the cut-off score even though a model is acceptable, for example in cases of small models, non-normality and the sample size, which is why we did not improve model fit by the use of modification indices (Marsh et al., 2004; Saris et al., 2009).

For sensitivity analyses, we repeated the hierarchical structural equation analyses with the non-imputed data and without the use of a robust estimator.

We computed power for the genetic covariance models using the shiny app "twin power calculator" (https://shiny.cnsgenomics.com/TwinPower/,Visscher, 2004). With 141 MZ twin pairs and 116 DZ twin pairs, power to detect A was .78 or above when A was at least .40 (see Table S2.3). Power for the linear multiple regression in the complete sample was excellent (.99) using G*Power (Faul et al., 2007) with an medium effect size of f = .15 (corresponding to β = .15), alpha of .05, 8 predictors and a sample size of 257. Using a sample size of 128 (randomly dividing the child sample), power was still good (.88) with a medium effect size of f = .15 (corresponding to β = .15), alpha of .05, 8 predictors. With a sample size of 257, alpha of .05, 8 predictors and a power of .80, we were able to detect effect sizes of .06 and above.

Transparency and openness

We report how we determined our sample size, all data in-/exclusions, all manipulations, and all measures in the study. All data is available upon request. Analysis code and research materials are available at the Open Science Framework platform. This study's design and its analysis were not pre-registered.

Results

Descriptives

2

Descriptive statistics, outliers and percentage of missingness of all variables can be found in Table 2.1. Table 2.2 shows correlations between the study variables.

Table 2.1

Descriptive statistics, outliers & missingness of study variables

Variable	N	*M*	*SD*	Range	% Missing
Parents					
Family SES	256	2.37	0.64	1.00 - 3.00	0.4
Wildman	248	2.23	0.62	1.00 - 4.25	3.5
PGS EA	136	0	1	-3.07 - 2.19	47.0
PGS Income	136	0	1	-2.67 - 2.38	47.0
PGS IQ	136	0	1	-2.68 - 2.11	47.0
Children					
IQ Total	512	103.58	11.76	72.5 - 137.5	0.4
Wildman	460	1.84	0.41	1.00 - 3.20	10.5
PGS EA	238	0	1	-3.19 - 2.57	53.7
PGS Income	238	0	1	-2.88 - 2.60	53.7
PGS IQ	238	0	1	-2.10 - 2.97	53.7

Note. SES = socioeconomic status; PGS = polygenic score; EA = educational attainment; IQ = general intelligence; N = sample size; M = Mean; SD = Standard deviation.

Table 2.2

Correlations of the study variables including confidence intervals

Variable	1	2	3	4	5	6
1. Family SES		.16	-.05	-.05	.05	.04
		[.07, .24]	[-.13, .04]	[-.13, .04]	[-.04, .13]	[-.04, .13]
2. Child IQ	.04		.04	.17	.26	.16
	[-.14, .21]		[-.05, .12]	[.08, .25]	[.17, .33]	[.08, .25]
3. Wildman	-.03	-.02		.00	-.24	-.06
	[-.15, .09]	[-.14, .11]		[-.09, .09]	[-.32, -.16]	[-.14, .03]
4. PGS EA	.28	.20	-.15		.60	.32
	[.16, .39]	[.08, .32]	[-.27, -.03]		[.54, .65]	[.24, .40]
5. PGS Income	.23	.11	-.22	.64		.46
	[.11, .34]	[-.01, .23]	[-.33, -.10]	[.56, .71]		[.39, .53]
6. PGS IQ	.02	.16	-.06	.31	.40	
	[-.10, .14]	[.03, .27]	[-.18, .07]	[.20, .42]	[.29, .50]	

Note. Upper triangle depicts children data and lower triangle depicts parent data. Child IQ and family SES are depicted only for Child 1 for readability. Bold estimates are significant at p,.05. SES= socioeconomic status; PGS= polygenic score; EA= educational attainment; IQ = general intelligence

Table 2.3

Cross-twin correlations and standardized estimates of the univariate ACE models

	MZ	DZ	A	C	E
Wildman parent score	.80	.74		.76	.24
Wildman child score	.52	.17	.44		.56
IQ Child Total	.55	.44		.50	.50
IQ Child Verbal	.43	.29	.43		.57
IQ Child Performance	.52	.42	.53		.47

Note. MZ= monozygotic;DZ= dizygotic; A = additive genetic component; C = common environmental component; E = unique environmental component; IQ = general intelligence.

Twin models

Cross-twin correlations and standardized estimates of the ACE models can be found in Table 2.3. For almost all variables, an AE model provided the best fit and goodness of fit indices can be found in Table S2.4. In the parents, a CE model provided the best fit for their responses to the Wildman questions, however, note that these estimates are computed in a child-based twin design (Kretschmer, 2023) in which the effects of parents' genetic makeup are represented in estimates of the C component in the ACE model. Heritability of the child-reported Wildman score was 44%. For the complete IQ scale, the best fitting model was a CE model, whereas in both subscales (performance and verbal IQ scale) an AE model provided better fit than a full ACE model. Heritability of the IQ subscales ranged between 43% and 53%. Unstandardized estimates including confidence intervals are reported in Table S2.5.

PGS Structural equation models

A structural equation model was employed to examine the sample. Full information maximum likelihood was used to include all data in the model estimation. Moreover, we utilized the Yuan-Bentler scaled Chi-square estimator with the Huber-White covariance adjustment to the standard errors of each parameter estimate (using "MLR" in lavaan). We created a latent variable (EDINQ) with three indicators (PGS of educational attainment, PGS of IQ and PGS of income) with good model fit (X^2 (3) = 5.61, p = .13, CFI = 0.98, TLI = 0.96, RMSEA = 0.07). Figure 2.1 shows the structural equation model for the parents and both children. For the hypothesized parent model of the Wildman score, model fit was good (X^2 (24) = 30.17, p < .18, CFI = 0.97, TLI = 0.95, RMSEA = 0.03). The model explained 8.6% of the variance in the outcome variable. The Wildman score was significantly predicted by the latent factor of the PGSs of EDINQ (β = -0.22, p = .002).

For the model with the first randomly chosen child, model fit was poor (X^2 (24) = 61.02, p < .001, CFI = 0.68, TLI = 0.55, RMSEA = 0.12). This can be expected in models with many variables and a modest sample size, and we therefore did not try to improve model fit using modification indices (Kenny & McCoach, 2003). The model explained 17.9% of the variance in the outcome variable. The Wildman score

was significantly predicted by the latent factor of the PGSs of EDINQ (β = -0.37, p < .001). For the model with the second randomly chosen child, model fit was also poor (X^2 (24) = 70.52, p < .001, CFI = 0.65, TLI = 0.51, RMSEA = 0.13). The model explained 21.0% of the variance in the outcome variable. The Wildman score was significantly predicted by the latent factor of the PGSs of EDINQ (β = -0.34, p = .014). All model estimates can be found in Table S2.6.

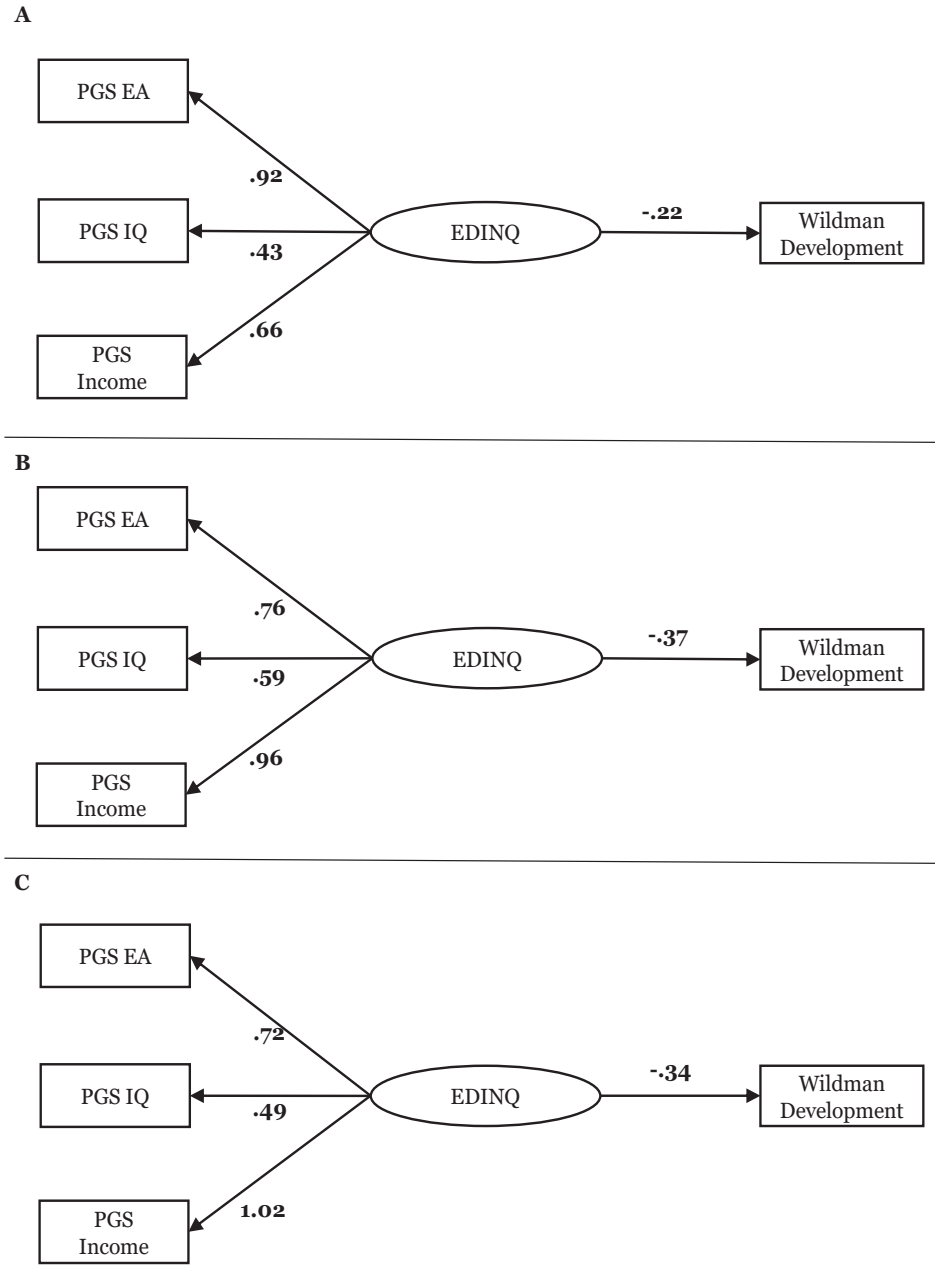
Sensitivity analyses

When conducting the analyses without a robust estimator, similar results emerged. The Wildman score was significantly predicted by the latent variable (β = -0.22, p = .001 in the parents; β = -0.37, p < .001 in child 1 and β = -0.34, p < .001 in child 2). In sensitivity analyses using only complete cases, the latent EDINQ factor of the PGSs only predicted Wildman scores in parents (β = -0.25, p = .026), and in the second child sample (β = -0.51, p < .001) but not the Wildman score in the first child sample (β = 0.19, p = .244).

We also investigated each of the PGSs that were previously combined in a latent factor separately. The Wildman score of the parents was significantly predicted by the PGS of educational attainment (β = -0.18 p = .002) and the PGS of income (β = -0.19, p = .002). The Wildman score of both child samples was predicted by the PGS of income (β = -0.33 p < .001 in the first child and β = -0.31, p = .005 in the second child) and in child 1 also predicted by the PGS of IQ (β = -0.18, p = .039). As additional analyses, we added the phenotype family SES to the parent model and the phenotype child IQ to the child model to examine whether the phenotypes would be associated with the Wildman score above and beyond the polygenic scores. The addition of either SES or IQ did not led to a significant improvement of the model in the parent and the first child ($X^2_{difference}$ (1) = 0.03, p = .86 in the parent model; $X^2_{difference}$ (1) = 0 55, p = .46 in the model of child 1), but it did in the second child ($X^2_{difference}$ (1) = 9.48, p = .002). On the other hand, the addition of the latent factor to a model with SES or IQ as a predictor led to a significant improvement of the model ($X^2_{difference}$ (1) = 9.46, p < .01 in the parent model; $X^2_{difference}$ (1) = 12.48, p < .001 in the model of child 1; $X^2_{difference}$ (1) = 8.79, p < .01 in the model of child 2). All estimates of the sensitivity analyses can be found in Tables S2.7 – S2.11.

Figure 2.1

Structural Equation model of PGSs predicting a) parents Wildman score and (b & c) children's Wildman score



Note. PGSs = polygenic scores; EA= educational attainment; IQ = general intelligence

Discussion

The acquiescence response bias in the Wildman Developmental Questionnaire showed considerable heritability. Parents as well as children displayed a genetic component in their responses to the absurd questions of the Wildman. First, using a twin design, we found that a considerable part of variance in the acquiescence response bias in the children was explained by an additive genetic component. Monozygotic twin pairs showed more similar answers than dizygotic pairs. The parents did not seem to answer the Wildman about each of their children differently depending on the zygosity of the twins. Whatever the twin zygosity, parents gave rather similar responses to the absurd Wildman items, which suggested a substantial common influence. Because of the child-based instead of parent-based twin design (Kretschmer, 2023) common influences we found despite diverging environments that dizygotic twin children create for their parents represent parents' own characteristics which are potentially influenced by their genotypes.

Secondly, we also found a substantial effect of molecular genetic predictors on the bias in the Wildman. A higher score on the latent factor of the moderately correlating polygenic scores of educational attainment, income and IQ predicted less bias in agreeing to impossible statements. Each of the three PGSs were also examined separately. The PGS for IQ did not influence the Wildman responses of parents or the children. The PGSs for educational achievement and income both affected the acquiescence response bias in the parents but in the children the PGS for educational attainment failed to predict their responses to the Wildman. Interestingly, the phenotypic scores for IQ and socio-economic status did not explain variance in the bias to agree to absurd statements although they were significantly associated with their corresponding PGSs.

Jerome Kagan was wrong in arguing that phenotypically measured socio-economic status would be more powerful in predicting adult or child behavior than most neurobiological assessments. We showed at least one exception to this rule: Polygenic scores for educational achievement or for income are better predictors of the acquiescence response bias than measured socio-economic status. One of the reasons might be that PGSs can be measured with less error than socio-economic status reported by the parents. We must however remind ourselves of the origin

of PGSs in GWAS that also had to use error-prone phenotypic assessments of the dependent variable, i.e. socioeconomic status or income. Another reason may be that PGSs cannot show reverse causality or a bidirectional association with dependent variables whereas such an alternative interpretation might always be possible in case of phenotypes that develop over time and under the influence of a myriad of other, unmeasured factors.

Surely Kagan was right at target in being vastly concerned about the frequent use of questionnaires in developmental research, and broader: in the social and behavioral sciences. We found evidence for genetic determinants of individual differences in the acquiescence response bias, not only on the level of behavior genetics but most importantly also with molecular genetic methods. Respondents are genetically predisposed to be more (or less) biased in their answers to absurd questions. Because widely used questionnaires like the SDQ present with several ambiguous, multifaceted items to which no response is adequate, this is a serious problem. Kagan (2007) emphasized in his content analysis of seemingly simple temperamental items such as 'My infant is afraid of strangers' stark ambiguity because they cannot be answered without being explicit about the meaning of terms like 'afraid' and 'stranger', about the context of the child's feelings, and about standards of comparison with other children or settings. The multiplicity of content triggers general, genetically anchored response biases to go along with what is perceived as the least compromising option.

In fact, even without any (understandable) question the acquiescence bias might emerge as a preference for specific middle-of-the-road answers. A genetic bias for this response bias is more problematic than just a tendency to answer in a socially acceptable way for two reasons. First, the genetic core of the bias cannot be influenced by exposure or outcome in developmental studies as the causal direction will be toward the environmental or developmental phenotypes instead of the other way around. This is the reason why polygenic score are preferred instrumental variables in causal research (Hamaker et al., 2020). The bias therefore will be a classic confounder inflating the associations between exposures and outcomes. Epigenetics is suggested to be shaped by the environment and in turn influences the expression of genotypes but some of our recent work suggests a more complicated

picture (Min et al., 2021). Second, in large scale genome-wide analysis studies on complex developmental traits such as temperament one would not want to run the risk of confusing an often rather tiny genetic signal for temperamental inhibition or reactivity with the genetic signal of a response bias. In GWAS research this risk should be addressed by routinely controlling for a polygenic score of the acquiescence response bias, comparable to the usual controls in GWAS for ethnicity to avoid population stratification. Developing a polygenic score for acquiescence response bias directly from a large scale GWAS project covering multiple cultures would be of great help to integrate such a score in any study using questionnaires.

Some limitations may be noted. Our genetic samples showed a high number of missing data which was due to asking for informed consent a few years after collecting DNA. Missing data lead to less power and in case of high numbers of missingness, imputation may be biased. However, our sensitivity analyses using complete cases yielded convergent results. A second limitation is the relatively homogeneous socio-economic status of the families who mostly came from a middle to higher socio-economic background. Future research should investigate more diverse families and examine (non-) linearity of the association between our PGSs and response bias. Replication is paramount as our results are based on a rather small sample. Lastly, we only addressed the acquiescence response bias and in future studies other response biases should be tested for a molecular genetic loading as well, such as the halo effect.

Nevertheless, if replicated, a possible genetic effect on response bias should be a red flag in times of replication crisis. One of the reasons for a replication crisis in psychology and related sciences is the use of too many researcher degrees of freedom (Simmons et al., 2011). In correlational cross-sectional or longitudinal studies, it is not uncommon to have respondents completing a variety of self-report questionnaires to measure (almost) the same constructs. This might easily lead to a fishing expedition in search for (just) significant results (p-hacking, Simmons et al., 2011) which turn out to be non-replicable, false positives. In large cohort studies this problem is exacerbated because of the so-called 'crud factor' (Meehl, 1990) meaning that every measured psychological variable is more or less weakly correlated with every other variable even if they are paired without theoretical basis merely through

random assignment (Meehl, 1990). The genetic loading we found for the acquiescence response bias might explain the existence of this crud factor because the underlying PGSs remain the same for each questionnaire as they are only dependent on the form of the questions and optional responses, and independent of the content of the items that should be assessed. This genetic crud factor may make null hypothesis testing obsolete because the null is non-existent.

For future research, assuming replication of the current findings, we need to overcome our questionnaire addiction and choose more sophisticated observational, ambulatory or interview measures and tests despite the larger investment of resources needed. Of course, going cold turkey on the questionnaire addiction is not realistic. Some intermediate steps are badly needed. Questionnaires might be preferred with items closer to observable behavior and by asking for frequency of specific child behaviors in the past week or so, much like Rothbart's (1981) Infant Behavior Questionnaire and its derivatives. Another requirement might be that independent informants should be asked to complete the same temperament questionnaire. The commonality between the various informant reports might be taken as the index for a temperamental characteristic. Third, temperament variables aggregated across several days in a week and across several settings should be preferred (Hamaker et al., 2020). Last, in their widely cited paper on the physiology and psychology of behavioral inhibition Kagan et al. (1987) promoted the idea of complementing psychological descriptions of temperamental inhibition with physiological assessments to create a firmer basis for this phenotype. These intermediate requirements for more valid measures of temperament may lead to a moratorium on studies with questionnaires completed by a single informant, on a single time-point and in a single setting and without a neurobiological component. This would be a great intermediate step, not only for research on temperament but also on other complex phenotypes such as attachment, executive functions, prosociality, externalizing or internalizing behavior problems.

Paradoxically, this is critically important for large genome wide association studies in search of polygenic scores as well. GWAS did teach developmental researchers some important lessons to overcome the replication crisis, with admirable features like striving for large consortia, pre-registration, built-in

replications and meta-analyses, and stringent corrections for multiple testing. But if GWAS rely on a few self-report questions to measure complex phenotypes, for example 'happiness' (Ward et al., 2022), a substantial genetic crud factor is almost guaranteed, and the resulting PGS for happiness might turn out to be in part an indicator of the acquiescence response bias. Large numbers of respondents and SNPs might not compensate for crudely measured phenotypes with high risk of response bias. Kagan (2007, p.369) would argue that this asymmetry in the sensitivity of the GWAS and questionnaire measures, that is, the disbalance between genotyping and phenotyping "is analogous to using an atomic clock to determine if a person is walking very slowly, slowly, moderately fast, or very fast".

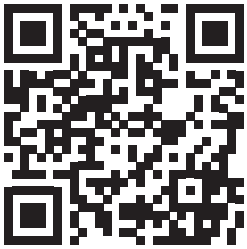
Acknowledgements

We thank the participating families for their enthusiastic involvement in the Leiden Consortium on Individual Development (L-CID). We are also grateful to the data-collection and data-processing team, including all current and former students, research assistants, PhD students and post-doctoral researchers for their dedicated and invaluable contributions. Marinus van IJzendoorn, Eveline Crone, and Marian Bakermans-Kranenburg designed the L-CID experimental cohort-sequential twin study "Samen Uniek" as part of the Consortium on Individual Development (CID; Gravitation Grant 2013-2023 awarded by the Dutch Ministry of Education, Culture, & Science, and the Netherlands Organization for Scientific Research, NWO Grant Number 024.001.003).

Supplementary Materials

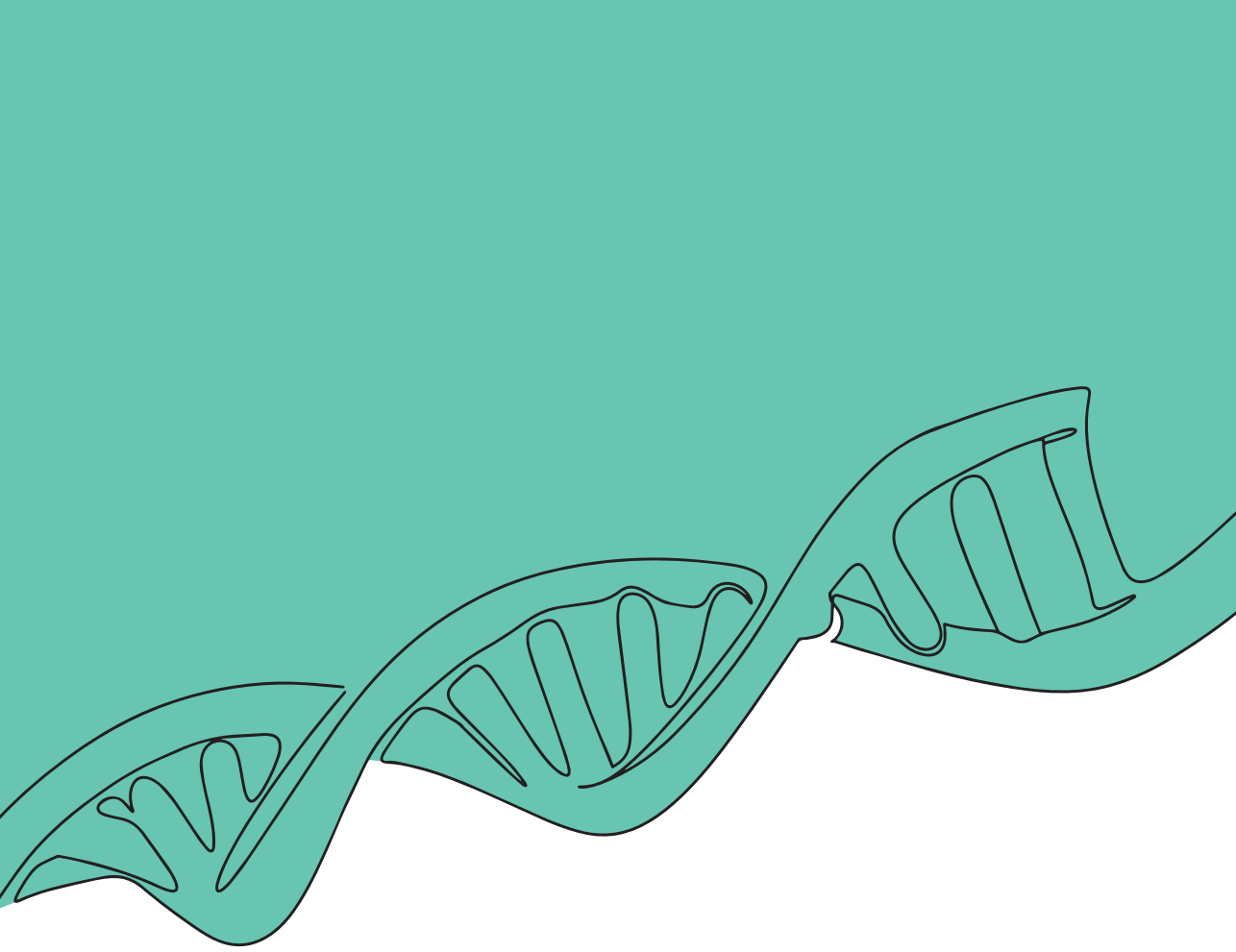
All supplementary materials are published and can be retrieved from:

http://tinyurl.com/Chapter2Supplement



This chapter is published as:

Runze, J., Bakermans-Kranenburg, M.J., Cecil, C.A.M., Van IJzendoorn, M.H., & Pappa, I. (2023). The Polygenic and Reactive Nature of Observed Parenting. *Genes, Brain, and Behavior, 22*.



THE POLYGENIC AND REACTIVE NATURE OF OBSERVED PARENTING

CHAPTER THREE

Abstract

In Wertz et al. (2019), parents' polygenic scores of educational attainment (PGS-EA) predicted parental sensitive reactions to the child, as observed in a dyadic task (i.e., observed sensitivity). We aimed to replicate and expand these findings by combining longitudinal data, child genotype data, and several polygenic scores in the Generation R Study. Mother-child dyads participated in two developmental periods, toddlerhood (14 months old; n = 648) and early childhood (3-4 years old, n = 613). Higher maternal PGS-EA scores predicted higher observed sensitivity in toddlerhood (*b* = .12, 95% CI .03, .20) and early childhood (*b* = .16, 95% CI .08, .24). Child PGS-EA was significantly associated with maternal sensitivity in early childhood (*b* = .11, 95% CI .02, .21), and the effect of maternal PGS-EA was no longer significant when correcting for child PGS-EA. A latent factor of PGSs based on educational attainment, intelligence (IQ), and income showed similar results. These polygenic scores might be associated with maternal cognitive and behavioral skills that help shape parenting. Maternal PGSs predicted observed sensitivity over and above the maternal phenotypes, showing an additional role for PGSs and yielding results that are more easily interpreted in a causal manner. In conclusion, we replicated the central finding of Wertz et al. (2019) that parental PGS-EA partially explains parental sensitivity. We found some evidence of evocative gene-environment correlation (rGE), emphasizing the dynamic nature of parenting behavior across time, although further research using family trios is needed before definitive conclusions can be drawn.

Parenting is a complex phenotype, ranging from sensitive responsiveness and limit-setting to harsh and neglecting approaches, and it is supposed to substantially shape child development. Previous studies on (mostly mothers') sensitive responses towards the child (i.e., maternal sensitivity) documented the predictive, positive associations with children's cognitive and language development, self-regulatory executive functioning, socio-emotional development, and fewer externalizing behavior problems (Borairi et al., 2021; Cooke et al., 2022; Pinquart, 2017; Valcan et al., 2018; Verhage et al., 2016). Antecedents of parenting such as parental socioeconomic status, mental health, and experiences of adversities have been found to explain part of the variance in sensitivity (Booth et al., 2018) but the direction of effects is not always clear. In this study, we use a polygenic score approach to investigate the role of educational attainment, intelligence, and income in predicting maternal sensitivity. These polygenic scores are assumed to refer to purely cognitive problem solving abilities (represented by PGS-IQ) as well as to noncognitive abilities implied by educational and social success (represented by PGSs for educational achievement and for income), such as planning skills, task persistence, or stress regulatory abilities (Heckman & Masterov, 2007). Cognitive and noncognitive abilities may both be relevant for effective parenting.

Previous (mostly child-based) twin studies on various dimensions of observed or self-reported parenting have shown moderate genetic effects, especially for reported parenting (Euser et al., 2020; Klahr & Burt, 2014). More similar parenting behavior in families with monozygotic versus dizygotic same-sex twins might be interpreted as evocative child effects on parental sensitive limit-setting (Euser et al., 2020), and a meta-analysis documented substantial evocative gene-environment correlations for parental positive reactions (e.g., structuring and sensitivity; Avinun & Knafo, 2014). However, molecular genetics studies examining candidate genes to explain variance in parenting showed equivocal results, and the main effects of (sets of) candidate genes seemed difficult to replicate (Mileva-Seitz et al., 2016). The introduction of polygenic scores (PGSs) based on genome-wide association studies (GWAS) incorporating numerous single nucleotide polymorphisms (SNPs) is a promising development in the area of parenting research because multiple genes are thought to be involved in the complex phenotype of parenting. Indeed, one study

showed evidence for SNP heritability (h2 SNP = .10, 95% CI .00, .19, N = 6,453) of self-reported parenting using genome-wide complex trait analysis (GCTA) (Culpin et al., 2020), a tool that estimates the variance explained by all SNPs instead of any particular SNP (Yang et al., 2011). Conducting a GCTA however requires large samples and even the ALSPAC cohort in which the study was performed was deemed to be underpowered for this approach (Culpin et al., 2020).

An alternative methodology is the application of PGSs derived from the published combined results of consortia with GWAS data on hundreds of thousands of participants. In the ground-breaking Dunedin study with observed parenting and parental GWAS data (Wertz et al., 2020), the authors used the PGS for educational attainment (PGS-EA) based on GWAS data of more than a million participants (23andMe Research Team et al., 2018) to predict variance in parenting in their sample of 702 participants. They found that parents with higher PGS-EA provided more warm, sensitive, and stimulating parenting to their 3-year-old children. Part of this association, however, might be evocative child effects as children inherit parental genes that might lead to children's phenotypical traits (e.g., aggression) that trigger specific parenting behavior (e.g., harsh limit setting). In the Dunedin study, the children's genomes were not assessed, and the authors tested for child effects indirectly, deriving temperament-like child traits from video-recorded child interactions with the parent. Based on this temperament measure, they found no evidence of evocative child effects. In a follow-up study, the authors used data from the E-Risk study, where observed parenting and genetic data of both parent and child were available. In that study, evocative gene-environment correlations between children's genetics and dimensions of parenting were found (Wertz et al., 2020). The divergent findings may result from different measures (observed child behavior versus genetic child data) or other factors (e.g., age of the child at the time of assessment). Additional studies are required to clarify the role of genetics in parenting and the extent of child genetic effects on parenting.

In the current study, we aimed to replicate the main finding of the Dunedin study, in particular the effect of the maternal PGS-EA on observed parenting, measured at two developmental time points, namely in toddlerhood (14 months) and early childhood (3 and 4 years). Maternal sensitivity at its core generally shows

continuity during development (Bornstein et al., 2017), but there is evidence that maternal sensitivity is also adaptive to changes in child development and may vary across time (Mills-Koonce et al., 2008), likely in response to the child's needs and behaviors. To better capture the dynamic nature of maternal sensitivity, we included two developmental periods that are characterized by rapid changes in a child's needs and challenge parenting behavior in a different way.

In addition, we controlled for evocative child PGS-EA effects extracted from the children's genomes. Next, we extended our search for polygenic effects on observed parenting by adding other relevant maternal PGSs, namely the PGSs of general intelligence (PGS-IQ) and income (PGS-income). Low maternal education, general intelligence and income have previously been associated with lower maternal sensitivity (Neuhauser, 2018; Van Doesum et al., 2007), while higher maternal education and income are correlated with more supportive and sensitive parenting (Unternaehrer et al., 2019). Supportive and sensitive parenting refers to interactions in which parents are aware of their child's emotional and physical needs and respond appropriately and consistently. Using the additional genetic indicators for intelligence and income to predict observed parenting we expected to better capture the complexity of parenting. These three PGSs of EA, IQ, and income are expected to depict effects of both cognitive and what have been collectively called "noncognitive skills" (Demange et al., 2021). The "noncognitive skills" refer to factors such as planning skills, task persistence or delay of gratification, and stress regulatory abilities, that are considered to be equally important as IQ in explaining academic and employment outcomes (Gutman & Schoon, 2013). This study focuses on the effects of the cognitive and noncognitive (or conative) domains on parenting. In our opinion, these domains are underrepresented in parenting literature, where the focus lies often on the affective domain (i.e., mood and emotions of the parents). Finally, we used the relevant maternal phenotypes (i.e., educational level, IQ, and income) to test for associations with observed parenting and investigated whether the PGSs had any additional predictive power over and above the maternal phenotypes.

In sum, we aimed to replicate and extend the Dunedin findings, using data from the Generation R Study, a population-based prospective cohort study based in The Netherlands in which observational parenting data were available at two

developmental time points, toddlerhood (14 months) and early childhood (3 and 4 years) (Kooijman et al., 2016). Our first aim was to test the associations of maternal PGS-EA with observed sensitive parenting. We also tested whether child genetics explained part of these associations. Our second aim was more exploratory, in that we investigated whether including parental PGSs of IQ and income were associated with observed parenting and had predictive power over and above maternal education level, IQ, and income.

Materials and Methods

Setting

The mothers and children in this study were participants of the Generation R Study, a population-based prospective cohort study based in Rotterdam, the Netherlands (Kooijman et al., 2016). Mothers with a delivery date between April 2002 and January 2006 were enrolled in the Generation R Study. The Medical Ethical Committee of the Erasmus Medical Centre approved the study protocol; data collection and ethical issues were described in detail elsewhere. The study was preregistered at https://doi.org/10.17605/OSF.IO/2EN8Y.

Study population

A subgroup of 1,247 women and their children were invited to our research center for observational assessments during infancy and toddlerhood. This group (Generation R Focus cohort) is of Dutch ethnic origin (Tiemeier et al., 2012). The mean age of the mothers in our sample was 31 years (*SD* = 4.49). Twenty-five mothers had twins: in these cases, one sibling of each twin pair was randomly selected for analyses. No siblings or other relatives participated in this study. 56% (n = 704) of mothers and children participated in laboratory observations to assess maternal sensitivity at age 14 months. 59% (n = 740) participated in laboratory observations at 3 years and home visits at 4 years. Our final sample included n = 648 mothers and n = 613 children (51% boys) with both genetic and observed maternal parenting data in toddlerhood and early childhood, respectively.

Non-response analyses showed that dyads included in the analyses did not differ from the excluded dyads on child sex, maternal educational level, andmaternal

sensitivity (Kok et al., 2013).

Measures

Observed maternal sensitivity. Maternal sensitivity was observed first during a lab visit at the child's age of 14 months during free play (Cents et al., 2014), and then during a lab visit at the age of 3 years and a home visit at age 4 years with two tasks: building a tower and etch-a-sketch (Kok et al., 2013). Maternal sensitivity was coded with satisfactory intercoder agreement from video recordings (free play ICC = .79; tower task ICC = .75; etch-a-sketch ICC = .79) (Cents et al., 2014). An overall score for maternal sensitivity in early childhood was computed by combining the 3- and 4-year measurements, as previously described (Kok et al., 2013).

Maternal education, income, and IQ. Information about maternal education was obtained by questionnaire during enrollment in the Generation R Study and categorized as follows: high (34.9%, university degree), mid-high (24.5%, higher vocational training), mid-low (25.9%, > 3 years general secondary school, intermediate vocational training), and low (14.7%, primary school, lower vocational training, intermediate general school, or 3 years of less general secondary school). Information about net household income (76.0% > €2200 per month) was obtained by postnatal questionnaires completed by both parents.

Maternal IQ was measured when the children were around their sixth birthday (Mean Age = 6.0 ± 0.3 years) at the Generation R research center. Maternal non-verbal IQ was assessed using a computerized version of the Ravens Advanced Progressive Matrices Test (APM), set I.15 (Ghassabian et al., 2014). The mean intelligence score was 100 (SD = 15) for the whole Generation R sample, as expected.

Genotyping and imputation. A detailed description of the Generation R Biobank has been published (Jaddoe et al., 2007). Maternal blood samples were available for 1,247 mothers of the Generation R Focus cohort. All mothers were of European ancestry, confirmed using principal components analysis on GWAS data. DNA was genotyped using the Infinium Global Screening Array with Multi-Disease drop-in (GSA-MD), version 2. Child blood samples were collected from cord blood at birth (Illumina 610K Quad Chip) or from venipuncture during a lab

visit at around 6 years (Illumina 660K Quad Chip). The Illumina 610K and 660K were merged based on their overlapping SNPs. Only children of European ancestry were selected for further analyses. For both mothers and children, quality control was performed in PLINK (version 1.9) (Purcell et al., 2007), as previously described (Lamballais et al., 2021; Medina-Gomez et al., 2015). Briefly, SNPs were removed if the minor allele frequency was < 1%, the Hardy–Weinberg equilibrium (HWE) p-value was < 1e−6, or the SNP call rate was < 98%. Individual data were removed in cases of genetic and sex mismatch, excess rates of homozygosity of the genotypes (> 4 SD), and genotype quality (> 5% missing). After genotyping, a two-step genotype imputation was applied for both mothers and children using the 1000 Genomes Project (phase III release version 5), build GRCh37/hg19 as reference panel. Monomorphic SNPs (with MAF < 0.5%) and SNPs with low imputation quality (R^2 < 0.3) were excluded, resulting in 49,008,248 SNPs.

Polygenic score (PGS) approach. For the replication part of this study, we used the publicly available GWAS summary statistics (N = 766,345 individuals) based on the recent study of Okbay et al. (2022) to estimate maternal PGS-EA in our sample. This sample is overlapping with the previous study of (23andMe Research Team et al., 2018). The sample including the participants from 23andMe is not (yet) publicly available (Powell, 2021). For the extension part of this study, we used the GWAS catalog (Buniello et al., 2019) to find relevant GWAS with publicly available summary data. For the PGS-general intelligence (PGS-IQ) we used the study of Savage et al. (2018), based on N = 269,867 individuals. For the PGS-household income (PGS-income) we used the study of Hill et al. (2019), based on N = 505,541 individuals. See Supplementary Table S3.1 for more details. Maternal and child PGSs were estimated using two different methods, to investigate whether the choice of the method would influence the findings. For our main analyses, we use a PC+T (p-value based clumping and thresholding) method similar to Wertz et al., using the PRSice software (Choi & O'Reilly, 2019) to estimate PGSs. For all three PGSs, the summary statistics served as the base sample, and Generation R was the target sample. The mothers and children participating in our study were never included in the base dataset. For the PGSs only autosomal SNPs were used. PGSs were calculated using clump r^2 = 0.1, 250kb

at different p-value thresholds (i.e., 0.001, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, and 1). We chose to use the best p-value threshold approach for each PGS to explain the most variance using the largest R^2 and increase predictive power for subsequent analyses (see in Supplementary Table S3.2 the explained variance per threshold). Since the optimal p-value threshold depends on various factors, such as the effect size distribution, the power of the base and target data, the genetic architecture of the trait, and the fraction of causal variants, and is thus not known a priori, this process of selecting the best p-value threshold is important and comparable to tuning parameter optimization (Choi & O'Reilly, 2019). The risk of overfitting is minimal when a large number of SNPs is used for each threshold, as has been previously discussed (Krapohl et al., 2018). However, to replicate the initial study of (Wertz et al., 2019), we have also estimated PGSs using all SNPs (p = 1), see Supplementary Table S3.4.

For comparison, we used LDpred2-auto, a version of LDpred2 that does not require a tuning sample. LDpred2 uses the same GWAS summary statistics as previously mentioned and Linkage Disequilibrium (LD) information from an external LD reference sample to infer the posterior mean effect size of each SNP (Privé et al., 2021).

Statistical analyses

Outliers, i.e., data points deviating 3.29 *SD* or more from the mean, were winsorized. Maternal sensitivity and PGSs were standardized. We tested two structural models using Structural Equation Modeling (SEM) analyses to handle missing data and simultaneously estimate the effects on multiple outcomes. Models were adjusted for child sex and the first ten principal components of genetic ancestry, to further control for hidden population stratification. The first model (the Replication model) examined the association of maternal PGS-EA with observed maternal sensitivity, in two developmental periods. We additionally included children's PGS-EA to control for possible child evocative effects. The second model (the EDINQ model) extended the replication model and examined the combined role of the strongly correlated maternal PGS-EA, PGS-IQ, and PGS-income. For this model as well, we added child PGSs in a second step to test for evocative gene-environment correlations. For all

SEM analyses, we used the lavaan statistical package (Rosseel, 2012). We used full information maximum likelihood (FIML). The Yuan-Bentler scaled Chi-square estimator with Huber-White covariance adjustment to the standard errors of each parameter estimate was used for non-normally distributed data. Bootstrapping was used to obtain bias-corrected confidence intervals. Model fit was assessed with the Comparative Fit Index (CFI, Bentler, 1990), the Tucker–Lewis Index (TLI, Tucker & Lewis, 1973), and the Root Mean Square Error of Approximation (RMSEA, Hu & Bentler, 1998). Good model fit was assumed with CFI and TLI values greater than .95 and RMSEA smaller than .08 (Xia & Yang, 2019). All analyses were conducted using R, version 4.04 (R Core Team, 2017). Finally, we used linear regression to predict maternal sensitivity from maternal phenotypical education, income, and IQ, and to examine the association of maternal PGSs over and above the effect of the relevant maternal phenotypes in a two-step regression model (with maternal phenotypes added as predictors of maternal sensitivity in step 1, and maternal PGSs added in step 2). These analyses were performed in SPSS 28.0 for Windows (SPSS Inc, Chicago, IL).

Results

Sample characteristics

Descriptive statistics and bivariate correlations between observed maternal sensitivity at two time points and maternal and child PGSs are shown in Table 3.1. Maternal and child PGSs were related, as expected since children receive half of their genetic variants from each parent (e.g., maternal PGS-EA and child PGS-EA, $r = .52$). The somewhat elevated genetic correlation might point to some assortative mating in our sample, which has been previously indicated for cognitive abilities (Jackson et al., 2022; Torvik et al., 2022). Before conducting the main analyses, we checked whether age of the mother and sex of the child were significant covariates to include in our SEM models, but this was only the case for child sex (see Supplementary Table S3.2).

Table 3.1

Descriptive statistics of and bivariate correlations between the study variables

Variable	N	M	SD	Min-Max	1	2	3	4	5	6	7	8	9
1 Age mother (at intake, in years)	1247	30.80	4.49	16.51 – 46.34									
2 Maternal IQ	1138	100	15.00	55.00 – 120.00	**.21**								
3 Maternal sensitivity (t)	704	0.01	0.83	-3.84 – 1.88	.03	**.23**							
4 Maternal sensitivity (eac)	740	0.48	0.77	-1.76 – 2.96	.04	**.28**	**.15**						
5 Maternal PGS-EA	1247	0.00	1.00	-3.46 – 2.78	**.27**	**.38**	**.11**	**.15**					
6 Maternal PGS-IQ	1247	0.00	1.00	-5.03 – 2.29	**.27**	**.34**	**.08**	**.15**	**.58**				
7 Maternal PGS-Income	1247	0.00	1.00	-2.87 – 3.15	**.19**	**.29**	**.10**	**.14**	**.68**	**.41**			
8 Child PGS-EA	1247	0.35	0.82	-3.90 – 2.73	**.24**	**.32**	**.10**	**.19**	**.52**	**.50**	**.37**		
9 Child PGS-IQ	1247	0.43	0.68	-3.68 – 1.98	**.22**	**.31**	**.08**	**.17**	**.44**	**.73**	**.39**	**.64**	
10 Child PGS-Income	1247	0.31	0.97	-3.39 – 3.21	**.18**	**.29**	.02	**.12**	**.43**	**.36**	**.30**	**.72**	**.46**

Note. Significant correlations are shown in bold, t= toddlerhood, eac = early childhood, EA = educational attainment, IQ = general intelligence.

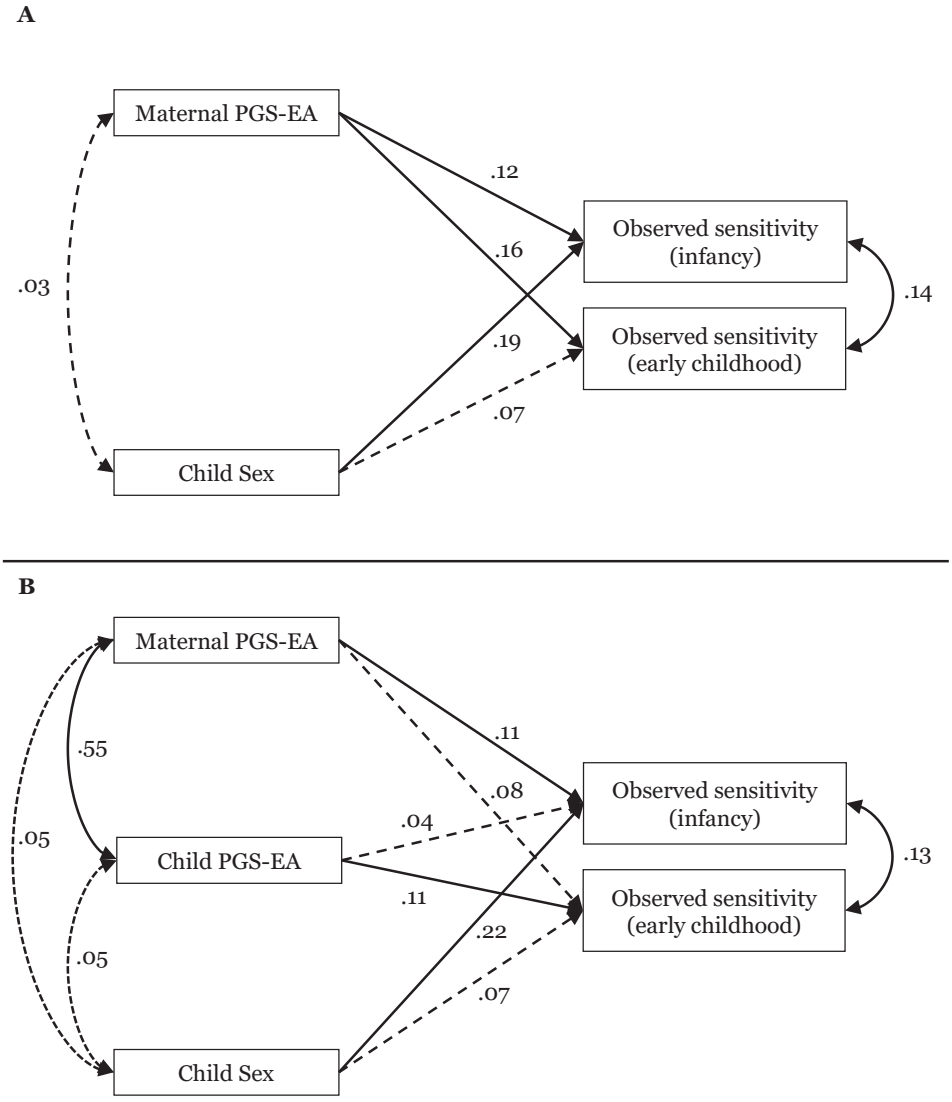
The Replication Model

Figure 3.1A displays the SEM results for the replication model. Parameter estimates and bootstrapped confidence intervals of the model are presented in Table 3.2. In this multivariate model higher scores on maternal PGS-EA predicted higher observed sensitivity in toddlerhood (*b* = .12, 95% CI .03, .20) and early childhood (*b* = .16, 95% CI .08, .24). The model explained 2.0% of the variance in observed sensitivity in toddlerhood and early childhood. Child sex (being a girl) was associated with more observed maternal sensitivity in toddlerhood (*b* = .19, 95% CI .04, .33), but not in early childhood.

We added the child's PGS-EA to control for the genetic effects of the child. Figure 3.1B shows the SEM results for this model, and Table 3.2 summarizes the estimates of the model. Child PGS-EA was not associated with maternal sensitivity in toddlerhood (*b* = .04, 95% CI -.05, .13), but it was significantly associated with maternal sensitivity in early childhood (*b* = .11, 95% CI .02, .21). Maternal PGS-EA was no longer significantly associated with sensitive parenting after accounting for child PGS-EA in early childhood (*b* = .08, 95% CI -.02, .18). R^2 increased from 2% in the maternal PGS-EA-only model to 3% in the model including the child PGS-EA. Sensitivity analyses using the maternal and child PGSs with a p-value of 1 showed similar results and are presented in Supplementary Table S3.5. Sensitivity analyses using LDpred2-auto confirmed that higher maternal PGS-EA is associated with more maternal sensitivity in both toddlerhood and early childhood (see Supplementary Table S3.5). However, child PGS-EA estimated using LDpred2-auto was not significantly associated with maternal sensitivity, in neither of the two time points.

Figure 3.1

A) Graphical representation of the replication model. This model tests the associations between the maternal polygenic score of educational attainment (PGS-EA) and observed parenting in two developmental periods (toddlerhood and early childhood). B) Graphical representation of the replication model controlling for child PGS-EA.



Note. single-headed arrows represent regression coefficients and double-headed arrows are correlation coefficients. Statistically significant estimates ($p < .05$) are shown with solid lines.

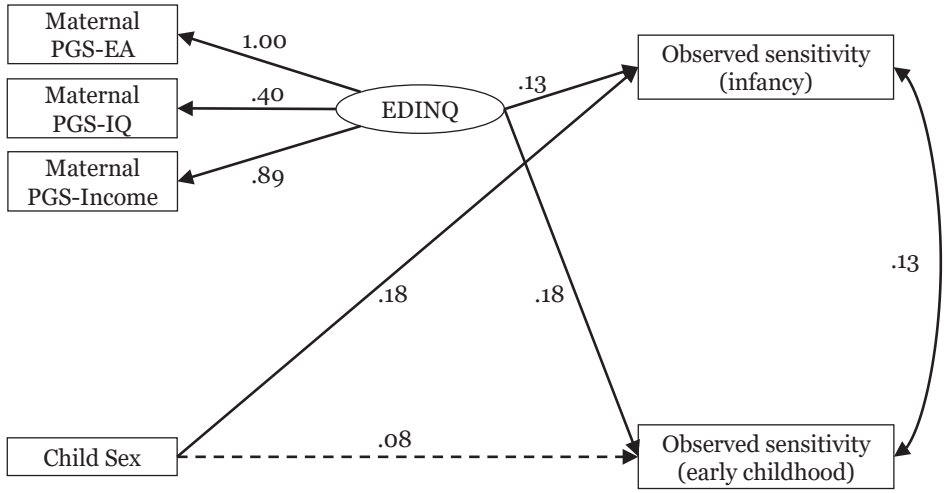
The EDINQ model

Figure 3.2 shows the SEM results for the EDINQ model, combining PGSs for educational attainment, income, and IQ. A latent factor was estimated by combining the highly correlated maternal PGS-EA, PGS-IQ, and PGS-income. Table 3.2 summarizes the parameter estimates. In this model, the latent factor was associated with maternal sensitivity in toddlerhood (b = .15, 95% CI .05, .26). Higher scores on the latent factor were also associated with higher maternal sensitivity in early childhood (b = .22, 95% CI .12, .31). The model explained 3% of the variance in observed sensitivity in early childhood. In supplementary analyses, we conducted multiple regressions with the three PGSs as independent predictors and results were comparable (see Supplementary Table S3.6).

Next, a latent EDINQ factor based on children's PGS-EA, PGS-IQ, and PGS-income was estimated and added to the model to investigate child genetic effects. For this model, the model fit was not acceptable (X^2 (22) = 158.94, p < .01, CFI = .949, TLI = .917). Therefore, based on modification indices, we added the correlation between the maternal PGS-IQ and the child PGS-IQ (mi = 89.08, epc = .13) which resulted in an acceptable model fit (X^2 (21) = 62.01, p < .01, CFI = .985, TLI = .974, RMSEA = .036). The latent factor of the child PGSs was not associated with maternal sensitivity in toddlerhood (b = .01, 95% CI -.13, .15). However, it was significantly associated with maternal sensitivity in early childhood (b = .16, 95% CI .01, .31), and the effect of the maternal latent factor was no longer significant (see Table 3.2 and Figure 3.3). R^2 increased from 3% in the maternal-only EDINQ model to 4% in the model including the child PGSs in early childhood. Sensitivity analyses using the maternal and child PGSs with a p-value of 1 showed similar results and are presented in Supplementary Table S3.5. Similar to the replication model, sensitivity analyses using LDpred2-auto confirmed that the latent factor of maternal PGS-EDINQ was associated with more maternal sensitivity in both toddlerhood and early childhood (see Supplementary Table S3.5). However, the child latent factor PGS-EDINQ estimated using LDpred2-auto was not significantly associated with maternal sensitivity, at neither of the two time points.

Figure 3.2

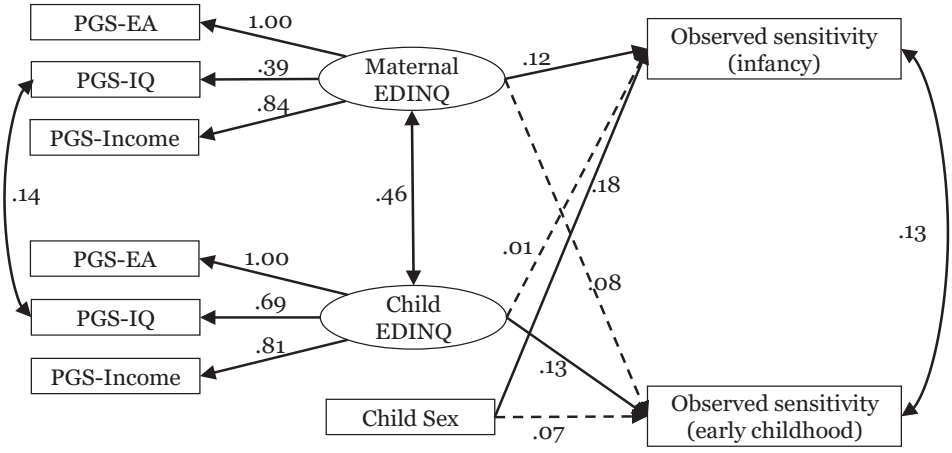
Graphical representation of the EDINQ model, combining maternal PGS-EA, PGS-IQ, and PGS-income in a latent factor.



Note. Observed variables are in rectangles and the latent variable is in a circle. Single-headed arrows represent regression coefficients. Statistically significant estimates ($p < .05$) are shown in solid lines.

Figure 3.3

Graphical representation of the EDINQ model controlling for child PGSs, estimated in a similar way as for mothers.



Note. Observed variables are in rectangles and the latent variable is in a circle. Single-headed arrows represent regression coefficients. Statistically significant estimates ($p < .05$) are shown in solid lines.

Table 3.2

Parameter estimates of the structural equation models

	Predictors	b	se	z	p	CI		R^2
Replication model: X^2 (0) = 00.00, CFI = 1.000, TLI = 1.000, RMSEA = .000								
Sensitivity (t)	Maternal PGS-EA	0.12	0.04	2.65	.008	0.03	0.20	.02
	Child sex	0.19	0.08	2.46	.014	0.04	0.33	
Sensitivity (eac)	Maternal PGS-EA	0.16	0.04	4.04	.000	0.08	0.24	.02
	Child sex	0.07	0.07	1.01	.314	-0.07	0.22	
Replication + child PGS-EA: X^2 (0) = 00.00, CFI = 1.00, TLI = 1.00, RMSEA = .000								
Sensitivity (t)	Maternal PGS-EA	0.11	0.05	2.22	.027	0.01	0.21	.03
	Child PGS-EA	0.04	0.05	0.92	.357	-0.05	0.13	
	Child sex	0.22	0.08	2.82	.005	0.07	0.38	
Sensitivity (eac)	Maternal PGS-EA	0.08	0.05	1.53	.126	-0.02	0.18	.03
	Child PGS-EA	0.11	0.05	2.27	.023	0.02	0.21	
	Child sex	0.07	0.07	0.94	.346	-0.08	0.21	
EDINQ model: X^2 (7) = 6.42, p = .49, CFI = 1.00, TLI = 1.00, RMSEA = .000								
EDINQ Mother	Maternal PGS-EA	1.00				1.00	1.00	.80
	Maternal PGS-IQ	0.40	0.03	15.43	.000	0.35	0.45	.25
	Maternal PGS-Income	0.89	0.05	18.63	.000	0.80	0.99	.55
Sensitivity (t)	EDINQ Mother	0.15	0.05	2.88	.004	0.05	0.26	.02
	Child sex	0.18	0.07	2.38	.017	0.03	0.32	
Sensitivity (eac)	EDINQ Mother	0.22	0.05	4.31	.000	0.12	0.31	.03
	Child sex	0.08	0.07	1.06	.290	-0.07	0.22	
EDINQ + child EDINQ: X^2 (21) = 62.01, p < .00, CFI = .985, TLI = .974, RMSEA = .036								
EDINQ Mother	Maternal PGS-EA	1.00	NA	NA	NA	1.00	1.00	.84
	Maternal PGS-IQ	0.39	0.02	16.30	.000	0.34	0.43	.25
	Maternal PGS-Income	0.84	0.04	22.96	.000	0.77	0.92	.52
EDINQ Child	Child PGS-EA	1.00	NA	NA	NA	1.00	1.00	.79
	Child PGS-IQ	0.69	0.04	17.42	.000	0.61	0.77	.34
	Child PGS-Income	0.81	0.04	20.19	.000	0.73	0.89	.79

Note. Nmax = 1,247, b = unstandardized parameter estimate, se = standard error, z = Z-statistic, CI = confidence interval, t= toddlerhood, eac = early childhood, mc = middle childhood; final model fits displayed in the table; *Model fit before modifications: X^2 (22) = 158.94, p < .01, CFI = .949, TLI = .917, RMSEA = .112; acceptable model fit was obtained after 1 modification: adding correlations between maternal PGS-IQ and child PGS-IQ (mi = 89.08, epc = 0.13); Both replication models are saturated, meaning that the number of free parameters is equal to the number of variances and unique covariances, which is why fit indices are not useful for these models.

Table 3.2 continued

Parameter estimates of the structural equation models

	Predictors	b	se	z	p	CI		R^2
Sensitivity (t)	EDINQ Mother	0.14	0.07	1.96	.049	0.00	0.28	.02
	EDINQ Child	0.01	0.07	0.15	.883	-0.13	0.15	
	Child sex	0.18	0.07	2.37	.018	0.03	0.32	
Sensitivity (eac)	EDINQ Mother	0.10	0.07	1.32	.187	-0.05	0.25	.04
	EDINQ Child	0.16	0.08	2.03	.043	0.01	0.31	
	Child sex	0.07	0.07	0.98	.328	-0.07	0.21	

Note. Nmax = 1,247, b = unstandardized parameter estimate, se = standard error, z = Z-statistic, CI = confidence interval, t= toddlerhood, eac = early childhood, mc = middle childhood; final model fits displayed in the table; *Model fit before modifications: X^2 (22) = 158.94, p < .01, CFI = .949, TLI = .917, RMSEA = .112; acceptable model fit was obtained after 1 modification: adding correlations between maternal PGS-IQ and child PGS-IQ (mi = 89.08, epc = 0.13); Both replication models are saturated, meaning that the number of free parameters is equal to the number of variances and unique covariances, which is why fit indices are not useful for these models.

Predictive power of PGSs controlling for maternal phenotypical education, income, and IQ

Table 3.3 presents the association between maternal phenotypes and maternal sensitivity in toddlerhood. Similar to the PGS-models, higher maternal education was associated with more maternal sensitivity in toddlerhood (β = .11, *p* = .01) and early childhood (β = .19, *p* = .001). In early childhood, higher household income was also associated with more maternal sensitivity (β = .10, *p* = .02). Maternal IQ was not associated with maternal sensitivity, neither in toddlerhood nor in early childhood.

Table 3.3 also presents the stepwise regression analyses for maternal sensitivity in toddlerhood, with maternal phenotypes added in step 1 and maternal PGS-EA and latent factor PGS-EDINQ added in step 2, respectively. Addition of maternal PGS-EA or PGS-EDINQ did not improve the toddlerhood model. For maternal sensitivity in early childhood, however, maternal PGS-EA and PGS-EDINQ increased the explained variance (adjusted R^2 increased from 7.5% to 8.0% and 8.4%, respectively), showing predictive power of maternal PGSs over and above the predictive role of the relevant phenotypes.

Table 3.3

Stepwise regression analyses testing the association between maternal phenotypes and observed maternal sensitivity in two developmental periods

		β	p	Adjusted R2	F
In toddlerhood/ replication model					
Step 1	Maternal education	**.11**	.01	2.2%	F(3,566) = 5.24, p =.001
	Household income	.07	.11		
	Maternal IQ	.05	.24		
Step 2	Maternal PGS-EA	.04	.35	2.2%	F(4, 565) = 4.15, p =.003
I*n toddlerhood/ EDINQ model*					
Step 1	Maternal education	**.11**	.01	2.2%	F(3,566) = 5.24, p = .001
	Household income	.07	.11		
	Maternal IQ	.05	.24		
Step 2	Maternal PGS-EDINQ	.06	.19	2.3%	F(4, 565) = 4.36, p = .002
In early childhood / replication model					
Step 1	Maternal education	**.19**	.001	7.5%	F(3,613) = 17.74, p <.001
	Household income	**.10**	.02		
	Maternal IQ	.08	.06		
Step 2	Maternal PGS-EA	**.09**	.04	8.0%	F(4, 612) = 14.46, p <.001
In early childhood/ EDINQ model					
Step 1	Maternal education	**.19**	.001	7.5%	F(3,615) = 17.93, p <.001
	Household income	**.10**	.02		
	Maternal IQ	.08	.06		
Step 2	Maternal PGS-EDINQ	**.10**	.01	8.4%	F(4, 614) = 15.19, p < .001

Note. EA = educational attainment; Significant effects are shown in bold.

Discussion

Although the heritability of parenting has been examined in twin studies in the past decades, Wertz and colleagues (2019) were among the first to use molecular genetics to investigate genetic effects in observed parenting. In the current study, we replicated their central finding of an association between mothers' polygenic score of educational attainment (PGS-EA) and sensitive interactions with their children. We found that already in toddlerhood a higher PGS-EA score was associated with higher sensitivity, suggesting that genetic differences shape phenotypic differences in parenting behavior at an early stage. The association between maternal PGS-EA and observed parenting further implies that cognitive (i.e., problem solving) and conative (i.e., planning skills, task persistence or delay of gratification, and stress regulatory abilities) processes may play an important role in shaping parenting, although the exact mechanisms are as yet unknown. It is important to note however that cognitive and noncognitive skills are often interlinked, and our ability to differentiate between them is limited.

An important contribution of our study is that we showed that differences in parenting are partly explained by genetic differences between children, as previously indicated (Dobewall et al., 2019; Shewark et al., 2021; Wertz et al., 2020). The association between maternal PGS-EA and observed sensitivity in early childhood was nullified when controlling for child PGS-EA. The inclusion of child genetic effects in the model increased the explained variance from 2% to 3%, highlighting an important path between child genotype and parental behavior. As proposed in Belsky's (1984) process model of parenting, child influences on parent-child interactions should not be neglected. Interestingly, we did not find child genetic effects on maternal sensitivity when measured in toddlerhood. Presumably, child temperament might only begin to exert effects on maternal sensitivity after toddlerhood (Plomin et al., 2016).

Our findings have several implications. First, parental sensitivity has mainly been considered as a parental trait (Ainsworth et al., 1974), predictable from pre-birth and more or less independent of child factors (Fonagy, 2022). Highly sensitive parenting is thought to compensate for difficult-to-handle features of the child, for example, temperamental irritability or reactivity (Mileva-Seitz et al., 2016). However,

moderately sensitive parents may respond more sensitively to easy-going children than to irritable children that require more patience in searching for the right response when distressed. The child's temperament or other features might create or increase a gap between parental competence and parenting performance.

Second, the association of the child's PGS-EA with maternal sensitivity at the expense of the mother's PGS-EA provides some evidence of evocative gene-environment correlation (rGE) (Knafo & Jaffee, 2013). Parental sensitive interaction is essentially dyadic and a two-way traffic of information, signals, and emotions, with parents in the lead but children as active participants. This study used mother-child genetic and phenotypic data. Since father genotype data were unavailable, our findings should be interpreted with caution. The observed direct genetic effects of the child may actually also include unmeasured paternal genetic effects on the parenting environment. To test whether our rGE interpretation is valid, larger studies with genetic data of family trios (child, mother, father) and observed parenting are needed (Harden & Koellinger, 2020; Knafo & Jaffee, 2013).

We extended the original model of (Wertz et al., 2019, 2020) by including other relevant PGSs in the cognitive and conative domain. Given strong correlations between PGSs for EA, IQ, and income we aggregated the three PGSs into a genetic indicator, PGS-EDINQ. The advantage of such higher-order aggregate might be better reliability and broader (ecological) validity, in particular when we compute a PGS-EA or PGS-EDINQ for offspring. Indeed, higher PGS-EDINQ scores predicted more observed sensitivity in early childhood, with an effect size similar to that of PGS-EA. The inclusion of child PGS-EDINQ decreased the effect of maternal PGS-EDINQ similarly as in the model with maternal and child PGS-EA. Conceptually PGS-EDINQ makes more sense in explaining parental sensitivity, because it provides a broader index of the context of parenting, although the aggregated factor did not predict substantially more variance. The loadings of maternal PGS-EA and PGS-income on the latent construct seem larger than the loading of PGS-IQ. The substantial association of this latent construct (EDINQ) with observed parenting may imply that cognitive (i.e., IQ) and conative (i.e., planning skills, task persistence or delay of gratification, and stress regulatory abilities) processes both play an important role in shaping parenting. But the purely cognitive problem-solving abilities

(PGS-IQ) might play a somewhat smaller role than the noncognitive or conative components. The exact mechanisms remain however still uncharted, maybe also because cognitive and noncognitive skills are interlinked and impossible to sharply differentiate even at the genetic level.

Another contribution of this study is the inclusion of and control for relevant maternal phenotypes (i.e., maternal education, household income, and maternal IQ) as predictors of maternal sensitivity. Higher income and IQ predicted more observed sensitive parenting, replicating earlier research (Neuhauser, 2018). Yet, maternal PGS-EA and PGS-EDINQ increased explained variance, over and above the related phenotypes, emphasizing the role of PGSs as a valuable tool in family studies.

The current study has some limitations. First, it is limited in statistical power because of the relatively modest number of participants. However, the study was pre-registered and replicated a previously published study. The replication part is thus transparent and reproducible without much leeway for researcher degrees of freedom. Furthermore, this study was based on the Generation R Focus cohort, which included observed measures of maternal sensitivity in only mothers of European ancestry. This homogeneity increases statistical power, but also implies limitations to the generalizability of the results and the need to broaden this line of research to other ancestries. Second, genotypes were imputed to the 1000 Genomes reference panel. The Haplotype Reference Consortium (HRC) reference panel is larger and might be preferred in samples of European ancestry such as ours. However, for replication purposes we used a methodology as similar as possible to the original Dunedin study. Third, in this study we used a strictly statistical approach to combine the highly correlated PGSs of EA, IQ and household income, conceptually similarly to previous work (Neumann et al., 2022; Warrier et al., 2021), and we added a sensitivity analysis with the three PGSs as separate predictors showing converging results. Other approaches, such as using genomic SEM to estimate a common factor among highly correlated traits and then creating PGS of the common factor, would focus on the joint genetic architecture of these traits, and eventually increase statistical power (Grotzinger et al., 2019). In this study, we focused on PGSs of the broader cognitive and conative domain. Based on Belsky's model of parenting (Belsky, 1984), PGSs of personality traits and psychopathology could also be of interest.

Although PGSs of the broader cognitive and conative domain have been found to predict parental sensitivity, we emphasize that this does not imply that parents' genetic make-up is defining their parenting skills. First, the prediction of parenting by polygenic scores is weak, especially compared to the prediction of parenting by numerous other factors, such as self-control skills of parents (Wertz et al., 2019). Polygenic scores might shape phenotypical traits and behaviors that in their turn predict parenting. Second, despite the association between parents' genetic make-up and their parenting behavior, we have shown that interventions can improve parenting behavior without altering genes, by changing the traits, circumstances or behaviors that mediate genetic influences on parenting (van IJzendoorn et al., 2023).

In sum, we replicated the Dunedin study on the relation between the polygenic score for educational attainment and observed sensitivity but also showed that the children's genotypic make-up has to be taken into account. Our results point to the role of evocative gene-environment correlation in the dynamic interactions between parents and children. Future studies could explore potential influences of other individual differences on parenting, such as (genetic) differences in personality and susceptibility to mental health problems. For an integral model of parenting larger and more powerful cohort studies are needed.

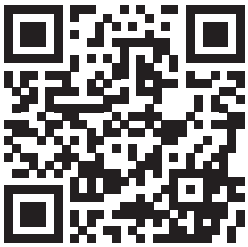
Acknowledgements

This study was supported by the Leiden Consortium on Individual Develop ment (L-CID)as part of the Consortium on Individual Development (CID; Gravitation Grant 2013-2023 awarded by the Dutch Ministry of Education, Culture, & Science, and the Netherlands Organization for Scientific Research, NWO Grant Number 024.001.003). The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsen-laboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam.

Supplementary Materials

All supplementary materials are published and can be retrieved from:

http://tinyurl.com/Chapter3Supplement



This chapter is published as:

Runze, J., Euser, S., Oosterman, M., Dolan, C. V., Koopman-Verhoeff, M. E., & Bakermans-Kranenburg, M. J. (2021). Actigraphic sleep and cortisol in middle childhood: A multivariate behavioral genetics model. *Comprehensive Psychoneuroendocrinology, 8,* 100094. https://doi.org/10.1016/j.cpnec.2021.100094

ACTIGRAPHIC SLEEP AND CORTISOL IN MIDDLE CHILDHOOD: A MULTIVARIATE BEHAVIORAL GENETICS MODEL

CHAPTER FOUR

Abstract

To date, behavioral genetic studies investigated either sleep or cortisol levels in middle childhood, but not both simultaneously. Therefore, a pertinent question is the degree to which genetic factors and environmental factor contribute to the correlation between sleep and cortisol levels. To address this question, we employed the classical twin design. We measured sleep in 6-9-year-old twins (N = 436 twin pairs, "Together Unique" study) over four consecutive nights using actigraphy, and we measured morning cortisol on two consecutive days. Sleep duration, sleep efficiency, and wake episodes were used as indicators of sleep. Morning cortisol level was used as cortisol indicator. A structural equation model was fitted to estimate the contribution of additive genetic effects (A), shared (common) environmental effects, (C) and unique environmental effects (E) to phenotypic variances and covariances. Age, cohort, and sex were included as covariates. The heritability of sleep duration, sleep efficiency and wake episodes were 52%, 45%, and 55%, respectively. Common environmental factors played no significant role. High genetic correlations between sleep duration and sleep efficiency and high genetic correlations between sleep efficiency and wake episodes were found. Shared environmental (29%), and unique environmental factors (53%) explained the variance in morning cortisol levels. Because the sleep and cortisol measures were found to be uncorrelated, we did not consider genetic and environmental contributions to the association between the sleep and cortisol measures. Our findings indicate that sleep duration, sleep efficiency and wake episodes in children are mostly impacted by genetic factors and by unique environmental factors (including measurement error).

Sleep and cortisol both have an impact on the psychological and physiological functioning of children (Bartels, De Geus, et al., 2003; Hatzinger et al., 2013; Sadeh et al., 2014). Sleep is controlled by two systems. There is the suprachiasmatic nucleus, which is the biological clock indicating that it is time to sleep when it is dark and time to be awake when it is light (Moore, 2007), and there is the sleep homeostat which keeps track of how much sleep we had and when we must refill the sleep reserve (Borbély & Achermann, 1999). The production of steroid hormone cortisol is regulated by the hypothalamus-pituitary-adrenal axis. Cortisol has several functions, including regulation of the metabolism, the immune system, and the response to stress (Oakley & Cidlowski, 2013). Its release from the HPA axis follows the circadian rhythm: It rises after midnight and rapidly at wake-up, peaking around half an hour after wake-up. Cortisol levels gradually decline during the day, with the lowest point at around midnight (Kirschbaum & Hellhammer, 1994).

Previous studies have suggested that sleep and cortisol are associated in children. Using a one-night measurement of polysomnography in 6-12-year-old children, Fernandez-Mendoza et al. (2014) found that children with short sleep duration and parent-reported sleep problems had increased evening and morning cortisol levels. Again using data based on one night of sleep, Lemola et al. (2015) also found a negative association between morning cortisol and sleep duration in 6-10-year-old children. Two studies have employed actigraphy to measure sleep in middle childhood. In a group of 282 8-year-old children, children with short sleep duration, compared to children with average sleep duration, had a higher cortisol awakening response, while children with low sleep efficiency, compared to children with average or high sleep efficiency had higher diurnal cortisol secretion (Räikkönen et al., 2010). El-Sheikh et al. (2008) found that higher levels of cortisol were associated with shorter sleep duration and poor sleep quality in a group of 64 children with a mean age of 8.75 years. Overall, shorter sleep duration has been related to higher cortisol secretion. However, most studies did not report correlations or effect sizes which is why we have no information about the strength of the association.

An open question concerns the extent to which genetic and environmental factors are involved in the association of sleep and cortisol. The physiological processes of sleep and cortisol are likely to be related. Cortisol-releasing-hormones

(CRH) have been suggested to be associated with sleep and wakefulness (Buckley & Schatzberg, 2005). High CRH has been associated with decreased slow wave sleep, and increased light sleep and wakefulness (Vázquez-Palacios et al., 2001). Suppression of CRH depends on inhibitory feedback by cortisol (Müller & Wurst, 2004). Glucocorticoid receptor (GR) activation at the level of the paraventricular nucleus decreases release of CRH, whereas GR activation on the amygdala increases release of CRH (Reul & Holsboer, 2002). Studies administering corticosterone have also reported either increased or decreased wakefulness or slow wave sleep (Buckley & Schatzberg, 2005; Vázquez-Palacios et al., 2001). Specifically, low doses seem to decrease wakefulness and increase slow wave sleep whereas high doses seem to do the opposite (Vázquez-Palacios et al., 2001). A possible explanation lies in the binding of cortisol to receptors. At low levels, cortisol binds to the MR, at higher levels cortisol binds to the GR and at very high levels, it binds to the amygdala GR (Buckley & Schatzberg, 2005). Binding to the amygdala GR is associated with an increase of CRH (and of subsequent cortisol) (Reul & Holsboer, 2002). Environmental factors can lead to high cortisol levels and/or sleep deprivation (Koopman-Verhoeff et al., 2019; Oakley & Cidlowski, 2013). Besides environmental factors explaining associations between sleep and cortisol, genetic factors may (partly) explain this association. Alexander et al. (2010) found a link between the brain-derived neurotrophic factor gene polymorphism (BDNF Val66Met) and stress reactivity: carriers of the Met allele had significantly less activation of the HPA axis in response to a psychosocial stressor. This might indicate that carriers of this polymorphism do not display high cortisol levels in which cortisol is binding to the amygdala GR impacting sleep.

To date, most twin studies investigating sleep have been conducted in adults, and a majority of these used self-reports of sleep duration and/or quality. Heritability coefficients of different sleep measures ranged from 30% to 60% (Gehrman et al., 2019; Genderson et al., 2013; Hublin et al., 2013; Lopez-Minguez et al., 2017). Behavioral genetic studies of sleep in children are scarce. Heritability coefficients for self- or parent-reported sleep quality or sleep duration range between 36% and 41% (Breitenstein et al., 2018; Taylor et al., 2015). A recent study investigated the heritability of some objective sleep measures, and the covariance between these measures in a sample of 8-year-old children (Breitenstein et al., 2021). Genetic

factors accounted for 81% of the variance in sleep duration, and for 79% of the variance in sleep efficiency. Common environmental factors played no role. This study also investigated the contributions of genetic and environmental factors to the covariance between sleep duration and sleep efficiency: A common genetic factor (genetic correlation of .85) underlies sleep duration and sleep efficiency. However, not all sleep measures were subject to this common genetic factor: e.g. sleep duration and sleep midpoint time variability did not share a common genetic factor. In this study, not all sleep measures were investigated together in a multivariate ACE model, which leaves the question open if different aspects of sleep, such as wake episodes may be subject to the same genetic and environmental factors. Different processes may be at play: pulsatile release of cortisol occurs when a person wakes up during the night. However, it is not clear whether these pulses build up to higher cortisol during the day (Buckley & Schatzberg, 2005). Given the low sample sizes in most actigraphy studies, a modest contribution of shared environmental factors may have gone undetected due to low statistical power.

Another important consideration is whether the extent to which genetic and/or environmental factors explain variation in sleep depends on the structure of the day of the children. A recent study in 11-to 14-year-olds indicated that common environmental factors do play a role on several sleep measures on school days, but not on free days, while genetic factors impacted sleep measures on the free days (Inderkum & Tarokh, 2018). It is not yet clear whether this pattern also presents in middle childhood samples, as adolescence is especially linked to changes in sleep, and no studies so far have investigated younger children using actigraphy.

Twin studies on cortisol levels have included samples ranging from 9-year old children to adults. In a study on 14-year-olds, heritability of the cortisol level after morning awakening was 28% (Ouellet-Morin et al., 2016). In a study on 9 to 16-year old children, genetic factors accounted for 28% of the variation in cortisol levels directly after morning awakening, 60% 30 minutes after morning awakening and 8% in the evening (Gustafsson et al., 2011).

Twin studies to date have considered either sleep or cortisol levels, but not both together. Therefore, little is known about the contributions of genetic and environmental factors to the covariance of sleep and cortisol. In the current study,

we employed a multivariate model based on the classical twin design to investigate the contribution of genetic and environmental factors to the phenotypic variance in actigraphy, cortisol levels, and to their phenotypic covariance. This study has three aims: 1) to investigate the contributions of genetic, shared environmental, and non-shared environmental factors to the variance in sleep duration, sleep efficiency, and wake episodes; 2) to investigate the contributions of genetic, shared environmental, and non-shared environmental factors to the variance in morning cortisol levels and the total cortisol production during the day; and 3) to investigate the contributions of genetic, shared environmental and non-shared environmental factors to the covariances of sleep and cortisol.

Based on previous research in adults (Genderson et al., 2013; Hublin et al., 2013), we expected genetic influences on sleep as well as cortisol. Sleep studies in adults did not report shared environmental influences of sleep measures. In our middle-childhood-aged sample, we expected shared environmental influences as we anticipated potential influences of school hours and parental rules about bedtime, which are shared by the twins, to affect sleep. Regarding cortisol, we expected genetic influences on the cortisol measures, but not shared environmental influences (Bartels, Van den Berg, et al., 2003; Steptoe et al., 2009; van Keulen et al., 2020). Lastly, we examined common genetic influences to explain the covariance between the cortisol and sleep measures, as previous research has found genetic influences on both cortisol and sleep (Bartels, De Geus, et al., 2003; Sletten et al., 2013). In exploratory analyses, we investigated whether the contribution of genetic and environmental factors differed depending on the day, testing for differences between weekdays and weekend days.

Methods

Participants

The "Together Unique" project is a longitudinal twin study with an experimental cohort-sequential design (see website: http://www.samen-uniek.com/) which includes two cohorts: an early childhood cohort (N = 239 families) and a middle childhood cohort (N = 257 families). Through municipal records, twin families from the western region of the Netherlands were invited to participate.

Families were contacted if the twins had the same gender, their parents were Dutch speaking, and both parents and grandparents were of European descent. Exclusion criteria were the presence of congenital disability, psychological disorder, chronic illness, hereditary disease, visual/hearing impairment, or an IQ of <70. For a detailed description of the recruitment, see Euser et al. (2016) for the early childhood cohort, and van der Meulen et al. (2018) for the middle childhood cohort. The study was approved by the central committee on research involving human subjects in the Netherlands (CCMO; Early childhood cohort NL49069.000.14; Middle childhood cohort NL50277.058.14).

The current study was pre-registered (https://osf.io/karqx/) and utilized data from the first measurement wave of the middle childhood cohort (collected in 2015/2016) (N = 256) and the fifth measurement wave of the early childhood cohort (collected in 2019) (N = 180), resulting in a total of 436 twin pairs (872 participants) of the same age range (6-9 years old). *N*=60 families dropped out of the study before the current fifth wave of the early childhood cohort. The mean age of the total group was 7.5 (*SD* = 0.59; early childhood cohort: *M* = 7.53, *SD* = 0.59, middle childhood cohort: *M* = 7.48, *SD* = 0.58), 48% were male, and 58% of the twins were monozygotic. Descriptives, separately for monozygotic (MZ) and dizygotic (DZ) twins, of both samples are shown in Table 4.1.

Procedure

All families received a set of actigraphs for their children via the mail. Actigraphs are watch-like devices which register sleep and wake states by recording movement. The children wore these devices on four consecutive nights, first two weeknights and then two weekend nights. All actigraphs were color-coded to reliably match them to each twin. During the research visit the parents were asked to download an app which served as an e-diary. Additionally, parents received daily/frequent reminders to use the actigraph. Parents were asked to put the actigraphs on their children's non-dominant hand wrist each night before the children went to sleep and to remove the actigraphs in the morning after the children woke up. Furthermore, parents received a paper logbook to report on any possible occurring problems regarding the app or actigraphs. Over the course of the two weekdays (first

two days of data collection), parents collected saliva samples from their children at three time-points (after awakening, between 16:00 and 18:00 and 30 minutes after dinner). Via the app, they were reminded to collect saliva and to report on whether the children had consumed any food or drinks, had engaged in physical activity, or had experienced any stress during the 30 minutes before saliva collection. The saliva was collected by requiring the children to spit into a small tube (passive drooling). Parents also reported in the paper logbook whether there were any issues during the saliva collection. Parental report on sleep times was used to check the data manually against errors in the measurement and/or analysis (see Measures). This information was used to enhance the accuracy of the actigraphy and e-diary data.

Measures

Sleep. Sleep was assessed for four subsequent nights, using wrist-actigraphy. Actigraphy is a well-validated and non-intrusive way to estimate sleep and wakefulness in the home (Sadeh, 2015). Previous research indicated that four nights are sufficient to reliably estimate the sleep measures included in this study (Acebo et al., 1999). Actigraphy has concordance rates of more than 90% in comparison with polysomnography, the gold standard for measuring sleep (Sadeh et al., 1994). MicroMini-Motionlogger actigraphs from Ambulatory Monitoring Inc. (Ardsley, NY) were used to collect data within fixed 1-minute time frames using Zero Crossing (ZC) mode. The sleep data was analyzed using Action-W software (Version 2.7.2305). First, we manually checked the actigraph data according to the Action-W user guidelines (Version 2.7.1). Bedtimes and rise times of the parental report were manually compared to the times of the actigraphs to detect possible inconsistencies in the actigraph data. Cases in which the parental report deviated 30 minutes or more from the actigraph data were reassessed independently by two raters (JR and MBK). In cases of rater disagreement, a third rater (MO) made the final decision. Subsequently, the following, commonly used, sleep variables were computed automatically using the validated Sadeh algorithm for children older than 12 months (Sadeh et al., 1989): sleep duration, sleep efficiency and wake episodes. Sleep duration is the total number of minutes asleep while being in bed. Sleep efficiency is the percentage of time being asleep between sleep onset

and morning awakening. Wake episodes are the number of occasions of adjacent 1-minute wake epochs while being in bed. We computed the means of sleep efficiency, sleep duration and wake episodes over the four days, as a recent study did not find substantial differences in sleep measures on weekend or on weekdays in this age group (Breitenstein et al., 2021). Nevertheless, we also computed the means for two weekdays and two weekend days separately to test this in our sample.

Cortisol. Saliva samples were taken three times per day over two consecutive days. Assessing cortisol via saliva has been widely used and proven as a reliable proxy for unbound cortisol in blood (Kirschbaum & Hellhammer, 1994). The non-invasiveness makes it easy to use in studies with children. By collecting saliva on two days, we obtained a more robust measure than with a single day, and we can account for non-compliance or mistakes during collection on one occasion. The primary parent performed the saliva collection directly after the awakening of the children. Parents received the instructions orally from the researcher beforehand as well as written in a paper logbook. Furthermore, parents received reminders via the app before each moment of collection ensuring that the time of saliva collection was adhered to. The primary parent performed the saliva collection 1) directly after the awakening of the children, 2) between 16.00 and 18.00 o'clock, and 3) 30 minutes after dinner. Due to the number of missing values on the second and third time point, and a high correlation between the morning cortisol and the daily cortisol production, we decided to deviate from our pre-registered plan and only use the morning cortisol samples in subsequent analyses. Parents were asked to store the saliva in the freezer. Once collected, saliva samples were stored at -20 degrees Celsius at the university until they were sent to the laboratory of the University Trier for cortisol analyses. To determine the cortisol concentration in the saliva sample, a time-resolved fluorescence immunoassay was used. On each batch, the same three saliva control samples (low, medium and high) were run. If control samples were out of a 2SD range, the whole batch was reanalyzed. The intra-assay coefficient of variation was between 4.0% and 6.7%, and the corresponding inter-assay coefficients of variation were between 7.1% -9.0% indicating that the variation between batches was low (Dressendörfer et al., 1992).

Inadmissible data values, as indicated by the lab, were treated as missing: These data values (in total six measurements or 0.1%) most likely did not reflect cortisol levels, but high glucose intake just before saliva sampling, which can distort the measurement accuracy of cortisol values. The mean morning cortisol levels of the two measurement days were used for the analyses. We expected the effect to be the strongest for the morning cortisol as cortisol builds up during the night (Kirschbaum & Hellhammer, 1994).

Zygosity. To determine the zygosity of the twins, DNA samples of the twins were taken by buccal swabs. For three twin-pairs in the early childhood cohort and one twin-pair in the middle childhood cohort, no DNA data was available. For these, zygosity was determined with a questionnaire filled in by the primary parent with eight items about physical resemblance and the confusion of the parents in distinguishing the twins (Rietveld et al., 2000). The questionnaire predicted zygosity in 93% of the cases compared to our own DNA analyses. Furthermore, two samples were re-analyzed as parents indicated their doubts concerning the zygosity determination. Most likely these two samples had been switched in the first analysis, as re-analysis showed that one twin pair was monozygotic instead of dizygotic and the other twin pair was dizygotic instead of monozygotic. These results were checked again and confirmed.

Covariates. To examine potential confounding factors influencing cortisol or sleep data, information on age, sex, and whether the twins shared a bedroom were collected using questionnaires, completed by the primary caregiver. Sleep duration and cortisol change during the development of children, therefore we controlled for age effects. Sharing a bedroom might also have an impact on sleeping patterns of the children. The data collection of the two cohorts was in different years, which also might have an impact on the data, so we investigated whether cohort was significantly associated with the phenotypes. Furthermore, in exploratory analyses, we checked whether sleep assessments took place in the vacation or during school times, in winter or summer time or during the change between winter and summer time.

Also, the following potential confounders of cortisol measures were included: body-mass-index, medication use, food or beverage intake 30 minutes prior to

sampling, physical activity 30 minutes prior to sampling or experience of stress 30 minutes prior to sampling (see Table S4.1 for frequencies of the dichotomous confounders). Studies indicated that obesity is associated with lower cortisol levels (Strahler et al., 2017). Medications can have an influence on salivary cortisol, such as corticosteroids, which act on the release of CRH or ADHD medication, where a dry mouth is a common side effect (Granger et al., 2009). Consumption of food or drinks, as well as physical activity and (psychological) stress have been found to increase unbound cortisol in saliva (Kirschbaum & Hellhammer, 1994; Strahler et al., 2017).

Statistical analyses

Data points deviating > 3.29 *SD* from the mean were winsorized in line with previous cortisol and sleep studies (e.g. El-Sheikh et al., 2008). Before conducting the main analyses, we computed descriptive statistics of the sleep and cortisol variables per cohort and for MZ and DZ twins separately. We also computed correlations and within-twin correlations for each variable (sleep efficiency, sleep duration, wake episodes and morning cortisol levels). ANOVAs and regression analyses were conducted with the potential confounding variables as predictors of the sleep and cortisol variables. Confounders with significant effects ($p < .05$) on the phenotypes were subsequently included in the main analyses. This resulted in the inclusion of only cohort.

We employed genetic covariance structure modeling as implemented in the OpenMx library in R (Boker et al., 2011) to estimate contributions of genetic and environmental factors to the phenotypic variances of, and covariances among, the sleep variables and cortisol levels using maximum likelihood estimation. The genetic covariance structure model included the contributions of additive genetic factors (A), shared environmental factors (C), and unshared environmental factors (E). The E factors may contribute to the phenotypic covariance among the four phenotypes (3 sleep phenotypes and cortisol), but do not contribute to the covariance between the twins (hence "unshared"). To estimate these contributions, we exploit the fact that MZ twins share 100% of their alleles (i.e., are genetically identical), while DZ twin on average share 50% of their alleles (Boomsma et al., 2002). Consequently, MZ

twins will show greater resemblance than DZ twin with respect to phenotypes that are subject to genetic influences. In covariance structure modeling, the MZ and DZ covariance matrices are modeled as follows:

		twin1	twin 2
S_{MZ} =	twin1	$S_A+S_C+S_E$	$S_A + S_C$
	twin2	$S_A + S_C$	$S_A+S_C+S_E$
S_{DZ} =	twin1	$S_A+S_C+S_E$	$.5* S_A + S_C$
	twin2	$.5* S_A + S_C$	$S_A+S_C+S_E$,

where S_{MZ} and S_{DZ} are the 8x8 phenotypic covariance matrices. S_A, S_C, and S_E are the 4x4 additive genetic (A), shared environmental (C), and unshared environmental (E) covariance matrices, respectively. The expected 4x4 phenotypic matrix equals $S_A+S_C+S_E$, the 4x4 twin 1-2 covariance matrices are S_A+S_C (MZs) and $.5*S_A+S_C$ (DZs). We first fitted the saturated model, in which we estimated the (unconstrained) MZ and DZ 8x8 covariance matrices. This saturated model served as a baseline model. Subsequently, we fitted the ACE model, where we parameterized the A, C, and E covariance matrices using the Cholesky or lower triangular decompositions. E.g., in the case of S_A, we parameterized S_A as $D_AD_A^t$, where D_A is a lower triangular 4x4 matrix (s1 to s3 are the sleep phenotypes), superscript t denotes matrix transposition:

		s1	s2	s3	cortisol
DA =	s1	a_{11}	0	0	0
	s2	a_{21}	a_{22}	0	0
	s3	a_{31}	a_{32}	a_{33}	0
	cortisol	a_{41}	a_{42}	a_{43}	a_{44}

The shared environmental and unshared environmental covariance matrices (S_C and S_E) were estimated in the same way: $S_C = D_CD_C^t$ and $S_E = D_ED_E^t$. In the ACE model, we tested the additive genetic, shared environmental and unshared environmental correlations between the sleep phenotypes and cortisol. In terms of the Cholesky parameterization, this test involves fixing parameter a_{41}, a_{42}, and a_{43} to

zero in the SA matrix, or analogous parameters in the Cholesky matrices of S_C (i.e., parameters c_{41}, c_{42}, and c_{43} in D_C) and S_E (i.e., parameters e_{41}, e_{42}, and e_{43} in D_E).

We conducted the statistical tests using the log-Likelihood Ratio Test (LRT) statistic. In this procedure, we fitted two models, models M0 and M1, where M1 is nested under M0, i.e., M1 can be derived from M0 by the imposition of parameter constraints. The test statistic used to evaluate the constraints is the minus twice the difference in the log-likelihood values of the two models. If the constraints are tenable, this (log-likelihood ratio) test statistic should follow a χ^2 distribution with the number of degrees of freedom equal to the difference in the number of parameters of model M0 and model M1. If the LRT statistic is larger than the critical value, associated with the given choice of alpha, the constraints are rejected. For instance, suppose model M0 includes SA based on the 10 parameters a_{11} to a_{44} (in DA), and model M1 includes SA but with the parameters a_{41}, a_{42}, and a_{43} fixed to zero. If these constraints are correct, then the LRT test statistic follows a chi-square (df = 3) distribution, in which 3 is the difference in the number of freely estimated parameters in M0 and M1. Given $\alpha = .05$, the critical value equals 7.81. That is, if the LRT test statistical is greater than 7.81, given df=3, we reject the constraints $a_{41} = a_{42} = a_{43} = 0$.

We adopted an alpha of .05 in carrying out the LRTs. In follow-up analyses, we assessed whether the contribution of genetic factors differed depending on the environmental structure of the day (weekdays versus weekend days).

Results

Preliminary analyses

The children slept on average 8.5 hours per night and had a sleep efficiency of 86%. The average number of 1-minute adjacent wake episodes was 24. The mean morning cortisol level was 9.35 nmol/l. All variables were approximately normally distributed (both skewness and kurtosis between -1 and 1). Variables containing outliers (deviating >3.29 *SD* from the mean) were winsorized: the 5% of the smallest and largest values were replaced by the lowest/highest retained values (see Table S4.2 for a list of frequencies of outliers per variable). There were missing data in all outcome variables. Of all children, 78% of the oldest twins provided the full four

nights of sleep measurement, 14% provided three nights, 5% provided two nights, 1.5% provided one night of measurement, and another 1.5% did not provide any sleep data. For the youngest twin, the numbers were 80% for four nights, 13% for three nights, 4% for two nights, 1% for one night, and 2% for no sleep data. The means and standard deviations of all mean sleep variables did not differ when we included only cases with complete sleep data or all cases with at least one night of sleep data (see Table S4.3). Therefore, there were 33 and 34 missing cases on each sleep variable for the oldest and the youngest twin respectively (8%). Missingness of the sleep variables is comparable to other child samples (9.8%) (Oullet-Morin et al., 2016). With regard to morning cortisol, 80% of the oldest (81% of the youngest) twin provided both morning measurements, 4% (5%) provided one morning measurement and 16% (14%) provided no morning cortisol data. There were 70 and 63 missing cases on the morning cortisol level for the oldest and the youngest twin, respectively (16% and 15%). We conducted Little's MCAR test to assess the missingness across the variables. Little's MCAR test was not significant (X (33) = 33.63, p = .44) indicating that missingness occurred completely at random. For our main analyses, the sample size was 402 twin pairs (235 MZ twin pairs and 167 dizygotic twin pairs). Sleep duration was positively correlated with sleep efficiency and negatively with wake episodes (p < .01). Sleep efficiency and wake episodes were also negatively correlated (p < .01). However, no sleep variables were significantly correlated to morning cortisol levels (MCL) (p > .05) (see Table 4.2). Therefore, we decided to exclude morning cortisol from the multivariate behavioral genetics model, as common genetic and environmental factors are most likely absent for unrelated variables. We report the results of the univariate twin model for morning cortisol.

We tested associations between possible covariates and the sleep and cortisol variables. 55% of twin pairs (59% of MZ twins and 51% of DZ twins) shared a room, however, room sharing was not associated with the sleep variables. For all three sleep variables, a main effect of cohort was found: sleep duration ($F(3, 846) = 13.17$, p < .001, adjusted R^2 = 4.1%), sleep efficiency ($F(3, 846) = 13.05$, p < .001, adjusted R^2 = 4.1%), and wake episodes ($F(3, 846) = 13.08$, p < .001, adjusted R^2 = 4.1%). All other variables (age, sex, BMI, medication use, drinking/eating, physical activity or stress 30 minutes prior to saliva sampling) did not have a significant association with the

sleep or cortisol variables. Therefore, only cohort was included in the twin analyses.

We computed the cross-twin within-trait correlations. The suitability of the ACE model can be evaluated by comparing the MZ and DZ correlations, where $2*r_{DZ}>r_{MZ}$ suggests that the ACE model is suitable. The correlations of the MZ twin pairs did not exceed twice the correlation of the DZ twin pairs for sleep duration and MCL. (Figure 4.1).

Multivariate ACE models

We fitted the full ACE model with the phenotypes sleep duration, sleep efficiency and wake episodes. The model fitted the data in comparison to the saturated model ($\Delta X^2[24] = 34.42$, $p = .078$; AIC = 5849.17 vs AIC = 5862.75). Table 4.3 shows the model fit statistics and Table 4.4 shows all parameter estimates including confidence intervals. Figure 4.2 displays the proportion of phenotypic variance and covariance accounted for by A, C, and E.

As can be seen in Figure 4.2, the heritabilities of sleep duration, sleep efficiency and wake episodes were 46%, 42%, and 55%, respectively. Unique environmental factors explained 47%, 55%, and 45% of the variance in sleep duration, sleep efficiency, and wake episodes, respectively. Common environmental factors played no significant role. A high genetic correlation was found between both sleep duration and sleep efficiency ($r = .91$) and sleep efficiency and wake episodes ($r = .71$).

Univariate ACE model for morning cortisol

We fitted the full ACE model for MCL ($\Delta X^2[3] = 3.53$, $p = .316$, AIC = 3956.83 vs AIC = 3959.29). The model fitted the data in comparison to the saturated model. Tables 4.3 and 4.4 show the model indices and heritability indices: 18% of the variance in MCL was attributed to genetic factors (note that the confidence interval included zero), 29% was explained by shared environmental factors, and 53% was explained by unique environmental factors.

Table 4.1

Means and Standard deviations of the Outcome Variables

	Total	MZ	DZ
	M (SD)	**M (SD)**	**M (SD)**
N twin pairs	436	251	185
Age	7.50 (0.59)	7.54 (0.61)	7.45 (0.55)
Sex % female	52	52	52
Bedroom % shared	55	59	51
Sleep duration in hours	8.54 (0.73)	8.51 (0.73)	8.56 (0.73)
Sleep efficiency in %	86.09 (6.19)	86.28 (6.16)	85.83 (6.23)
Wake episodes	24.10 (6.42)	23.78 (6.29)	24.55 (6.59)
MCL	9.35 (3.70)	9.40 (3.52)	9.28 (3.94)

Note. MCL = morning cortisol level in nmol/liter; means and standard deviations only of the oldest twin displayed for readability.

Table 4.2

Phenotypic correlations of the Outcome Variables

	1	**2**	**3**	**4**
1 Sleep duration in hours		.76* [.71, .79]	.31* [.22, .39]	.09 [-.01, .20]
2 Sleep efficiency in %	.80* [.76, .83]		.56* [.49, .63]	.08 [-.03, .18]
3 Wake episodes	.25* [.15, .34]	.49* [.41, .56]		-.03 [-.14, .07]
4 MCL	.06 [-.04, .17]	.04 [-.07, .14]	01 [-.09, .12]	

Note. MCL = morning cortisol level in nmol/liter; * p < .01, Correlations for the oldest twin are shown above the diagonal, for the youngest twin below the diagonal; 95% confidence intervals of the correlations are shown in brackets.

Table 4.3

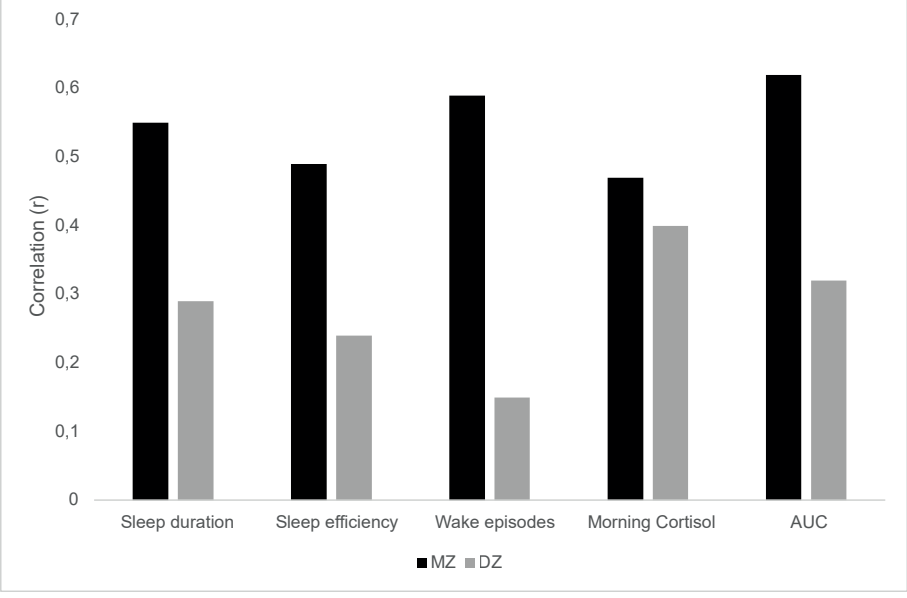
Full and best-fitting Cholesky decomposition fit statistics for the multivariate sleep model and for the univariate morning cortisol levels model

Model	Test	-2LL	df	AIC	Δ df	Δ χ2	*p*
Sleep Model							
0. Saturated Model		5766.75	2367	5862.75			
1. ACE-ACE-ACE	1 vs 0	5801.17	2391	5849.17	24	34.42	.078
MCL Model							
0. Saturated Model		3943.29	731	3959.29			
1. ACE	1 vs 0	3946.83	734	3956.83	3	3.53	.316

Note. The -2LL = -2 log-likelihood ratio test statistic; AIC = Akaike's information criterion; Δ df = change in degrees of freedom when model parameters were dropped; Δ χ2 = change in -2LL when model parameters were dropped; *p* = p-value of significance of the chi-square test; Cohort was included as a covariate.

Figure 4.1

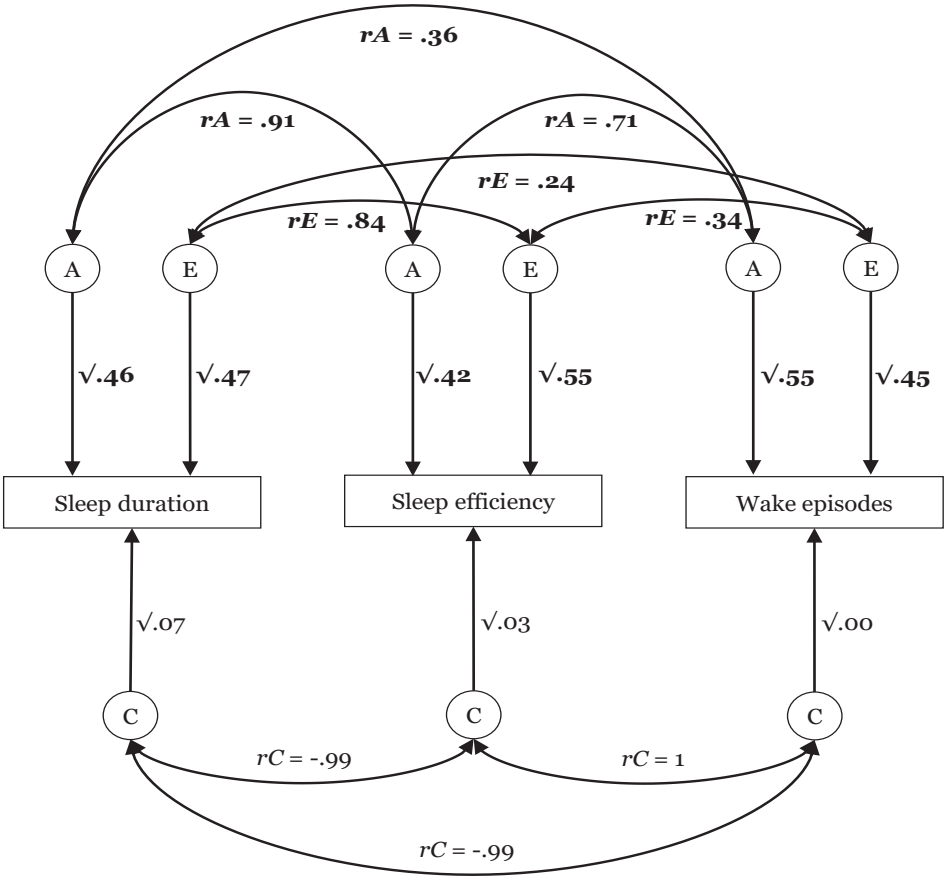
Cross-Twin Within-Trait Correlations



Note. MZ = monozygotic twin pairs, DZ = dizygotic twin pairs.

Figure 4.2

Final Multivariate Model for Sleep in Children



Note. The model with genetic and unique environmental effects for all variables and shared environmental effects for sleep duration; Variance components (i.e., squared standardized parameter estimates) shown for A, C, and E of all phenotypes; r_A is the genetic correlation between two variables, r_C is the shared environmental correlation between two variables, r_E = unique environmental correlation between two variables; Estimates with confidence intervals not including zero are shown in bold; Note that the C parameter of wake episodes is zero, which implies that the r_C correlations of the other two sleep phenotypes with episodes are not identified, and uninterpretable. Note that the other C parameters are also very low which renders the relevant r_C highly unreliable and therefore effectively uninterpretable; Confidence intervals omitted for readability (see Table 4.4); Cohort was included as a covariate.

Table 4.4

Full and best fitting Cholesky decomposition parameter estimates

	A	C	E
Multivariate Sleep model			
Sleep duration	.46 [.20 - .58]	.07 [.00 - .29]	.47 [.39 - .57]
Sleep efficiency	.42 [.29 - .52]	.03 [.00 - .13]	.55 [.46 - 65]
Wake episodes	.55 [.20 - .63]	.00 [.00 - .14]	.45 [.37 - .54]
	$\mathbf{r_A}$	$\mathbf{r_C}$	$\mathbf{r_E}$
Sleep duration x Sleep efficiency	.91 [.89 - .97]	-.99 [-1.00 - 1.00]	.84 [.81 - .87]
Sleep duration x Wake episodes	.36 [.06 - .48]	-.99 [-1.00 - 1.00]	.24 [.12 - .35]
Sleep efficiency x Wake episodes	.71 [.69 - .87]	1.00 [-1.00 - 1.00]	.34 [.23 - .44]
Univariate MCL model			
ACE	.18 [.00 - .51]	.29 [.01 - .50]	.53 [.43 - .64]

Note. r_A is the genetic correlation between two variables, r_C is the shared environmental correlation between two variables, r_E = unique environmental correlation between two variables; Cohort was included as a covariate.

Table 4.5

Means, Standard deviations and Pearson's Correlations for Weekdays and Weekend Days separately

		Weekday		Weekend day	
	Phenotype	**MZ**	**DZ**	**MZ**	**DZ**
M	Sleep duration	8.66 (0.77)	8.66 (0.74)	8.36 (0.83)	8.46 (0.89)
	Sleep efficiency	86.37 (6.42)	85.95 (6.57)	86.07 (6.94)	85.65 (7.18)
	Wake episodes	46.87 (18.33)	45.05 (17.43)	46.56 (18.49)	46.41 (24.17)
r	Sleep duration	.50	.20	.51	.35
	Sleep efficiency	.45	.19	.42	.27
	Wake episodes	.49	.09	.49	.07

Note. MZ = Monozygotic, DZ = Dizygotic; M = Mean; r = Correlation; means and standard deviations only of the oldest twin displayed for readability; Cross-twin-within-trait Pearson's correlations.

Exploratory analyses

As pre-registered exploratory analyses, we investigated the sleep variables on weekdays and weekend days separately. Because the sleep and the cortisol variables were unrelated, MCL was not included. The mean sleep duration, sleep efficiency and wake episodes did not differ much between weekdays and weekend days. Sleep duration was shorter on weekend days compared to weekdays (Table 4.5). Cross-twin-within-trait correlations also tended to be similar, with only DZ correlations on sleep duration and sleep efficiency being lower on weekdays compared to weekend days.

We ran the ACE model consisting of the three sleep variables for weekdays and weekend days separately. The ACE model for the weekdays showed adequate fit in comparison to the saturated model ($\Delta X^2[30] = 35.23$, $p = .234$; AIC = 19287.29 vs AIC = 19312.05). We tested an AE model but constraining the C parameters led to significant worsening of the model fit; therefore, the ACE model was retained. There was no common shared environmental factor explaining variance in sleep duration, sleep efficiency, and wake episodes. A high genetic correlation was found for sleep duration and sleep efficiency and a high genetic correlation was found for sleep efficiency and wake episodes. Heritability indices ranged between 33% and 48%, the unique environment accounted for between 52% and 66% of the variance in the phenotypes (Figure S4.2).

We also fitted an ACE model for the weekend days. This model fit poorly in comparison to the saturated model ($\Delta X^2[30] = 130.36$, $p < .001$; AIC = 19662.21 vs AIC = 19591.85). However, the explained variance and covariance in the phenotypes were similar to the weekday model. There was no common shared environmental factor explaining variance in sleep duration, sleep efficiency and wake episodes. A high genetic correlation was found for sleep duration and sleep efficiency and a high genetic correlation was found for sleep efficiency and wake episodes. Heritability indices ranged between 36% and 46%, the unique environment accounted for between 52% and 64% of the variance in the phenotypes (Figure S4.3). As measurement during holidays, winter-, or summertime and during the switch between these two were not associated with any of the sleep variables ($p > .05$), no further exploratory model was run with these variables as covariates.

Discussion

We performed a multivariate behavioral genetic study on sleep in school-aged children. We found that sleep duration, sleep efficiency, and wake episodes were moderately heritable. A high genetic correlation was found between sleep duration and sleep efficiency, and also between sleep efficiency and wake episodes. Shared and unique environmental factors played a role in the variance of morning cortisol levels. We did not investigate common genetic, shared environmental or unique environmental factors between actigraphic sleep and morning cortisol levels because there was no phenotypic correlation between actigraphic sleep and cortisol.

Heritability of sleep

Our findings with regard to the heritability of actigraphic sleep do not deviate much from the previous actigraphic sleep study in 12-year-old children (Sletten et al., 2013). Sleep duration was highly heritable in that study with 65%, which was comparable in our sample (46%). Sleep efficiency was also quite heritable with 52% in that study and 42% in the current study. Wake episodes had a heritability of 57% in the previous study and 55% in our study. The results of the small Sletten et al. (2013) study, including only 25 MZ twins and 41 DZ twins were thus replicated in our larger sample (251 MZ twins and 185 DZ twins). However, our heritability indices are somewhat lower in comparison to a more recent study (Breitenstein et al., 2021). Sleep duration was highly heritable with 81% in that study as well as sleep efficiency with 79%.

We also investigated whether the heritability of actigraphic sleep would be different on weekdays or weekend days. Contrary to Inderkum and Tarokh (2018), we found no differences. An explanation could be that our sample is much younger (6- to 9-year-olds compared to 11- to 14-year-olds). In adolescence, the need for sleep increases and adolescents have more freedom in deciding when to go to bed (Inderkum & Tarokh, 2018). In our sample, children probably go to bed at the same time every night independent of a week or weekend day which is also reflected in the reported waking/sleeping times. Moreover, in the Netherlands, sports competitions for child teams tend to be early in the morning on weekends, creating a more similar situation on a weekend day as compared to a weekday.

Genetic correlations in sleep variables

In the current study we found high genetic correlations between sleep duration and sleep efficiency and high genetic correlations between sleep efficiency and wake episodes, but low genetic correlations between sleep duration and wake episodes. Genome-wide association studies have predominantly focused on finding loci for sleep duration, and the PAX8 locus was consistently found to be associated with sleep duration (Dashti et al., 2019; Doherty et al., 2018). In a recent study, Dashti et al. (2019) detected 78 loci for self-reported sleep duration. These loci have also been found to associate with actigraphy derived sleep duration and sleep efficiency, indicating a common genetic background of both as identified in our study. Unfortunately, wake episodes were not included as a variable in these analyses. A GWAS on insomnia symptoms, in which participants were asked whether they would wake up often in the middle of the night, has identified several loci associated with insomnia symptoms (near MEIS1, TMEM132E, and CYCL1), but not with sleep duration in adults (Lane et al., 2017). This indicates that distinct genetic factors might be implicated in sleep duration compared to waking episodes. At the same time these genetic factors might as well be correlated, some of the loci associated with sleep duration and sleep efficiency might also be contributing to wake episodes. Therefore, more twin studies and more genome-sequencing research are needed to shed light on the contribution of different genes to the variation in sleep duration, sleep efficiency and wake episodes.

Heritability of cortisol

In this study, we found no significant heritability of morning cortisol levels. The literature shows a large variability in heritability estimates, depending on age, time of sampling, and type of measurement (urine, blood, or saliva). For example, Bartels, De Geus, et al. (2003) found in a sample of 12-year-old children that heritability of morning cortisol was 22%-24% when measured at 7.30am, but 56-59% when measured at 8.30am. In a meta-analysis the combined heritability for basal cortisol levels was 62%, but values ranged between 0 and 88% (Bartels, Van den Berg, et al., 2003). However, the review also indicated that power was insufficient for every singular study included in the review (Bartels, Van den Berg, et al., 2003).

Association between sleep and cortisol

We were surprised that there was no correlation between sleep and cortisol variables in our study. Previous research found significant associations between sleep and cortisol in middle childhood. However, there are indications that the association holds only for sleep problems, and not for the normal range of healthy sleep patterns in children: Fernandez-Mendoza et al. (2014) found that only the combination of reported insomnia symptoms with a shorter sleep duration was significantly related to morning and evening cortisol levels. Moreover, Pesonen et al. (2012) found that sleep problems were related to lower diurnal cortisol. In samples without reported sleep problems, sleep duration and sleep efficiency were related to afternoon cortisol levels in a healthy middle childhood sample (El-Sheikh et al., 2008). However, the mean sleep duration in that study was 6.58 hours, which deviates from the recommendation of nine to twelve hours for children of that age (Hirshkowitz et al., 2015). This is in line with another study that compared children with a normal sleep duration and children with a low sleep duration (less than 7.7 hours), and only found an association with cortisol for the short sleepers (Räikkönen et al., 2010). In line with our results, a recent study analyzing the association between sleep and hair cortisol in a sample of non-clinical children also found no association between sleep and cortisol (Eythorsdottir et al., 2020). Also, Marceau and colleagues (2019) investigated parent-reported sleep duration and morning cortisol in children longitudinally and did not find an association at any time point between 4.5 and 9 years.

Therefore, it might be that sleep and cortisol are not linearly related in the normal range of variation in sleep, but only in individuals with reported sleep problems or deviating sleep patterns.

Another possibility is that any association between sleep and cortisol may have gone undetected by looking at average sleep and cortisol parameters over several days. Several studies found that hours of sleep were associated with cortisol levels of the following day and cortisol levels in turn predicted the sleep hours of the following night (Van Lenten & Doane, 2016; Zeiders et al., 2011). This might indicate that sleep and cortisol may be associated in normal ranges of variation in sleep, but more subtle and subject to day-to-day fluctuations.

Strengths and limitations

To our knowledge, this is the first study combining sleep and cortisol in a behaviour genetics analysis using actigraphic sleep measurements over four days in the home environment and using cortisol measurements over two days. Furthermore, we are among the first to study heritability of cortisol and sleep in middle childhood. However, some limitations also must be noted. Although we tried to standardize the measurements as much as possible, some measurement error may be present. Having the parents put on the actigraphs is a great advantage when it comes to investigating sleep in a natural environment, but it also leads to less control over the measurements. The actigraphs might have been put on the dominant hand or put on the wrist too loose, affecting the measurements. The same is true for cortisol measurement, having the parents oversee the saliva collection in the home environment ensures that no external factors (nervousness in a laboratory setting for example) influence the measurement. At the same time, there is less control about whether the parents use the correct tubes and whether timing is in concordance with the instructions. Cortisol rises quickly in the morning after awakening, therefore timing of saliva collection is an important issue. In addition to instructing participants, we used an app to record the times that children woke up and that saliva was collected. However, parental reports of waking times were not convergent with actigraph data, and both measures may be biased in different directions: Children's early awakening may escape their parents' awareness, while actigraphy can be overly sensitive to waking and movement. Also, by studying cortisol and sleep in a children sample in their home environment, we could not prevent that some parents skipped some measurements, leading to missingness. Furthermore, some parents skipped the questions regarding possible covariates of morning cortisol (e.g. food intake before saliva collection). Although unlikely in the early morning, such unnoticed covariates might have affected the cortisol estimates. Thus, although the amount of missing data was similar to other studies (Marceau et al. 2019), future research should try to decrease these and other possible measurement errors by controlling the time of data collection more closely. An interesting avenue for the field of cortisol measurement are wearable devices that continuously measure cortisol. A first study by Parlak et al. (2018) has tested a device that detects cortisol in sweat. However,

more research and development are needed until such devices are suitable as an ambulatory assessment method in child research. Another point concerning the reliability of the cortisol data are possible batch effects, which were not taken into account in our analyses. However, quality control processes were employed in the laboratory. We recommend that future studies use a double control standard by also including the batch numbers in their analyses, which is already done in many other research fields.

The fact that the correlations between the sleep variables and MCL were found to be zero also raises the issue of power. Therefore, we conducted a power analysis to assess whether our sample size would have been sufficient to detect any the additive genetic, shared environmental and unshared environmental correlations between the sleep phenotypes and cortisol. Based on previous research, we expected that the phenotypic correlation between the sleep phenotypes and morning cortisol in children would be about .25. We explored the power to detect additive genetic correlations and unique environmental correlations of differing sizes. In the case of a genetic correlation of .2 and a unique environmental correlation of .4, the power to reject $a_{41}= a_{42}= a_{43}= 0$ (no contribution of A to the phenotypic correlations) was .38. Given genetic correlations of .6 and unique environmental correlations of .1, the power to reject $e_{41}= e_{42}= e_{43}= 0$ (no contribution of E to the phenotypic correlations) was .44, given an alpha of .05. In all other scenarios (genetic correlation of .8, .4, and 0), power was sufficient. We also did not find significant heritability estimates for the univariate twin model of MCL. Assuming a small A of .18, the power to detect A (given a C of .29 and E of .53) was .34.

Although we collected data from more than 400 twin pairs, resulting in more than 800 children engaged in four nights of ambulatory assessments, this sample size was still too small to have sufficient power to detect small additive genetic correlations or unique environmental correlations. Future studies should replicate our study with even larger samples to have a higher probability to detect possible correlations of A, C, and E.

Conclusion

In this study, we investigated the genetic factors implicated in the variation in actigraphic measured sleep and cortisol in children. Sleep is moderately heritable whereas cortisol levels are mostly explained by the shared and unique environment. High genetic correlations between sleep duration and efficiency were found as well as high genetic correlations between sleep efficiency and wake episodes. Sleep and cortisol were not related in our non-clinical low-risk sample. Future research should focus on disentangling the genetic contributions at play in aspects of sleep as well as investigating under which circumstances sleep and cortisol are correlated.

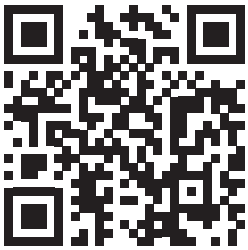
Acknowledgements

We thank the participating families for their enthusiastic involvement in the Leiden Consortium on Individual Development (L-CID). We are also grateful to the data-collection and data-processing team, including all current and former students, research assistants, PhD students and post-doctoral researchers for their dedicated and invaluable contributions. Marinus van IJzendoorn, Eveline Crone and Marian Bakermans-Kranenburg designed the L-CID experimental cohort-sequential twin study “Samen Uniek” as part of the Consortium on Individual Development (CID; Gravitation grant 2013-2023 awarded by the Dutch Ministry of Education, Culture, & Science, and the Netherlands Organization for Scientific Research, NWO grant number 024.001.003).

Supplementary Materials

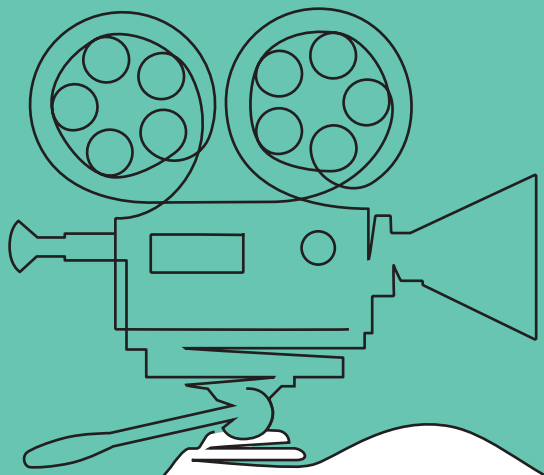
All supplementary materials are published and can be retrieved from:

http://tinyurl.com/Chapter4Supplement



This chapter is published as:

Runze, J., Van IJzendoorn, M. H., Vrijhof, C. I., & Bakermans-Kranenburg, M. J. (2022). Replicating a randomized trial with video-feedback to promote positive parenting in parents of school-aged twins. *Journal of Family Psychology, 36*(4), 490–501. https://doi.org/10.1037/fam0000961



REPLICATING A RANDOMIZED TRIAL WITH VIDEO-FEEDBACK TO PROMOTE POSITIVE PARENTING IN PARENTS OF SCHOOL-AGED TWINS

CHAPTER FIVE

Abstract

In this randomized controlled trial, we investigated the effectiveness of the brief, home-based Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD) in parents of 257 school-aged twin pairs (N = 514 children, Mean Age = 7.92, SD = 0.66), replicating a previous study testing the effectiveness of the intervention in parents with preschool-aged twins (Euser et al., 2021). We conducted two pretests (1 year apart) and one posttest 1 month after the intervention. An age-adequate twin-adapted version of the VIPP-SD was used in primary caregivers (91% female). We examined the main effect of the intervention on observed parental sensitivity and sensitive discipline and on attitudes toward sensitivity and sensitive discipline. We also investigated whether parents who are more susceptible to the environment, as measured by their self- and partner-reported current temperamental reactivity, benefitted more from the intervention. In our sample with older children, the VIPP-SD did not significantly change observed parental sensitivity or sensitive discipline in the intervention group compared to the control group. The VIPP-SD did improve parents' attitude toward sensitivity, but not toward discipline. Intervention effects were not moderated by temperamental reactivity of the parents, providing no support for the differential susceptibility hypothesis. Future research might examine the differential susceptibility hypothesis in parents using stress reactivity or genetic susceptibility markers instead of self-reported reactive temperament.

Despite the recent shift to more genetically informed explanations of variation in child developmental outcomes, it is crucial that researchers and practitioners also pay attention to parenting skills as they are consistently associated with child outcomes. Parental sensitivity has been related to positive child outcomes, such as social competence, better cognitive outcomes, less separation anxiety, and better emotion regulation (Dallaire & Weinraub, 2005; Frick et al., 2018; Leerkes et al., 2009; Martin et al., 2007). Harsh and insensitive parenting in turn have been associated with cognitive and behavioral problems in children, such as poor executive functioning, delinquency, and substance use problems (Deater-Deckard et al., 2012; Hinnant et al., 2015; Lucassen et al., 2015). Our aim was to examine how we can promote better parenting practices in parents of twins, as they might encounter more challenges in the upbringing of their children than parents of singletons. In this study we replicated an intervention study using the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD) with parents of preschooler twins (Euser et al., 2021) by testing its effectiveness in families with older, school-aged twin children.

From a clinical perspective, evidence-based interventions are much needed as negative child outcomes have a substantial societal impact. According to Heckman and Masterov (2007), early interventions have the highest return to human capital investment, making it crucial to focus on families with (young) children, instead of adults with established (behavioral) problems. However, solely focusing on preschoolers or even younger children, where the return of investment is the highest, might not be wise as many challenges for parents only emerge when their children grow older. In an individual participant data (IPD) meta-analysis, Gardner et al. (2019) did not find age-based differences in effectiveness of the Incredible Years parenting program in families with 2- to 12-year old children. These results indicate that also after the preschool period, parenting interventions can be successful in promoting child developmental outcomes and parent-child relationships.

Here we report on an effort to extend the age range of the VIPP-SD program to school-aged children. The VIPP-SD program is based on attachment theory and social learning theory and aims at enhancing parental sensitivity and sensitive discipline. The intervention is standardized and individualized, and has been tested

in several different samples: Clinical as well as nonclinical families, child care settings, and minority families (Juffer et al., 2017). In a previous study, the VIPP-SD has been used to support parents of 5- to 7-year-old twins (Euser et al., 2021). Families with twins are crucial for descriptive research on behavioral and molecular genetics of human development (Polderman et al., 2015), but they have often been "neglected" when it comes to supportive parenting interventions. However, raising twins may be even more challenging than raising one child. Challenges such as dividing the attention between the same-age twins and risks of jealousy and competition occur more frequently in twin families (Segal & Knafo-Noam, 2021). Furthermore, parents of twins have been found to be less sensitive compared to parents of singletons (Ellis-Davies et al., 2022).

In the twin sample of the Euser et al. (2021) study, the VIPP-SD was successful in increasing the level of parents' sensitive discipline. Differential susceptibility of parents to the intervention was also investigated, that is, whether parents with high temperamental reactivity would benefit more from the intervention than parents with low temperamental reactivity. No support for this hypothesis was found. The differential susceptibility hypothesis stems from the notion that some individuals might be more affected by environmental factors than others and that this susceptibility holds for positive as well as negative influences (Belsky et al., 2007). Although most research focused on children, studies also found support for differential susceptibility in adults (Slagt et al., 2015). In adults, mostly candidate genes (such as Dopamine Receptor D4;DRD4, Monoamine oxidase A; MAOA, or Serotonin transporter polymorphism; 5-HTTLPR) have been used as markers for susceptibility, but some support has been found for high sensitivity to the environment or openness to the environment as susceptibility markers (Hartman & Belsky, 2016; Slagt et al., 2015). Thus, it is important to take into account the differential susceptibility of the parents to not underestimate intervention effects (Bakermans-Kranenburg & van IJzendoorn, 2015).

Research Questions

The purpose of our preregistered study was to replicate the findings of the VIPP-SD randomized controlled trial with parents of preschooler twins (Euser et al.,

2021) in a sample of parents with school-aged twins. Our first research question was whether the VIPP-SD intervention would enhance sensitivity and sensitive discipline in parents of 9- to 11-year-old twins. In the preschooler twin study, no effect on parental sensitivity was found and we wondered whether the same would be the case in this older twin sample. Furthermore, according to the theory of planned behavior, any goal-directed behavior might be partly explained by the intention to show that behavior and this intention is in turn influenced by the attitude toward that specific behavior (Ajzen, 1991; Andrews et al., 2010; Hamilton et al., 2020). Therefore, we also investigated whether the VIPP-SD influenced attitudes toward sensitivity and sensitive discipline. Making use of the unique twin sample, we also examined whether intervention effects were different or similar for each twin sibling within a family, in particular in families with dizygotic (DZ) twins. Lastly, we tested whether parents were differentially susceptible to the intervention.

We hypothesized that (1) sensitivity and sensitive discipline of parents in the intervention condition would increase more or decrease less post-intervention, compared to parents in the control condition. Our secondary Hypothesis (2) was that positive attitudes toward sensitivity and sensitive discipline of parents in the intervention condition would increase more or decrease less post-intervention, compared to parents in the control condition. Furthermore, Hypothesis (3) was that the effects of VIPP-SD on parenting would be different for parenting behavior toward the two children within a family in which case the addition of the family level in the analytic model would explain variation in parenting. Our last Hypothesis (4) was that parents who are more temperamentally reactive would benefit more from the VIPP-SD than parents with lower reactivity.

Methods

Design

The “Together Unique” project is a longitudinal randomized controlled trial with families with twins (see website: http://www.samen-uniek.com/). Through municipality records, twin families from the western region of the Netherlands were selected. Families were contacted if the twins had the same sex, their parents

were able to communicate in Dutch and parents as well as grandparents were born in Europe. Families received an invitation letter and an information brochure. Parents who indicated a willingness to participate received a phone call to check the inclusion criteria and to provide more information about the study. The occurrence of congenital disability, psychological disorder, chronic illness, hereditary disease, visual/hearing impairment, or an Intelligent Quotient (IQ) of <70 led to exclusion. Six yearly visits were planned (alternating home and laboratory visits) with additional ambulatory assessments. The present study was pre-registered and utilized data from the first, second, and third assessment (T1, T2, T3; collected from 2015 to 2018). Approval for this study was provided by the central committee on research involving human subjects (Centrale Commissie Mensgebonden Onderzoek; CCMO; Middle childhood cohort NL50277.058.14). The study adheres to the CONSORT guidelines (see Appendix Figure A5.1).

Participants

One-thousand one-hundred and seventy four families were contacted through an invitation letter of which $N = 257$ families with twins were enrolled in the study ($N = 514$ children, 55% MZ pairs). Figure 5.1 shows the recruitment and randomization in a flow chart. At T1, the children were on average 7.92 years old ($SD = 0.66$) and 51.6% were female. The primary parents were on average 40.48 years old ($SD = 4.66$) and 91% were female. Fourteen families decided not to participate at T1, T2 or in the randomization process, resulting in a final sample size of $N = 243$. Table 5.1 shows the characteristics of the sample.

Intervention

Randomization. We randomized the sample at the family level in a ratio of 2:3, using a computer-generated blocked randomization sequence (block size = 19 families, stratified by timing of the intervention and twin sex). Due to limited resources, it was not possible to have a 50–50 split for intervention and control group, however, power of the study was only marginally affected by this ratio (power = .94 compared to .95 with a 50:50 split, with $\alpha = .05$ and assuming an effect of d = .42 as found in Juffer et al. (2018)).

We randomized the sample after T2 to minimize selective attrition. A researcher who was not involved in data collection or data coding assigned the families to one of both conditions using a random numbers generator. Ninety-one (37%) families were allocated to the intervention group and 152 (63%) families were allocated to the control group. Because of the open label design, interveners and participants were blind to the treatment condition before randomization but not afterwards. However, coders and researchers who were involved in data analysis were blind to the assignment.

Video-Feedback Intervention to Promote Positive Parenting and Sensitive Discipline for Twins. In the current randomized controlled trial (RCT), we implemented an age-appropriate version of the VIPP-SD adapted to twin families between T2 and T3 (see Figure 5.1) to promote positive parenting behavior. Adaptations to the original VIPP-SD program to address challenges faced by parents of twins were rather minor as previously described in Euser et al. (2016) and included the addition of tasks to discuss issues regarding dividing attention to both kids at the same time and dealing with jealousy and competition among the children. Also, toys and games were adapted to the situation of interacting with two children. In short, the VIPP-SD for twin families includes five biweekly sessions at the family's home conducted by an intervener. In our study, it was a female intervener, however this is not mandatory. At the start of each session, the intervener videotaped approximately 15 min of some standard situations to observe interactions between the children and the parent, such as playing a game together or performing a task. Next, video-feedback on the child's or parent's behavior and the interaction between them was provided, based on the recordings of the previous session. The feedback focused on positive and successful interaction moments and specific moments of effective positive parenting matching the theme of the session (see Juffer et al. (2017) for an overview of the themes). In order to support the primary parent's implementation of positive parenting behaviors, the partner of the parent was invited to join this booster session, according to protocol. For more information on the themes and methods of the intervention see Euser et al. (2021) and Juffer et al. (2017).

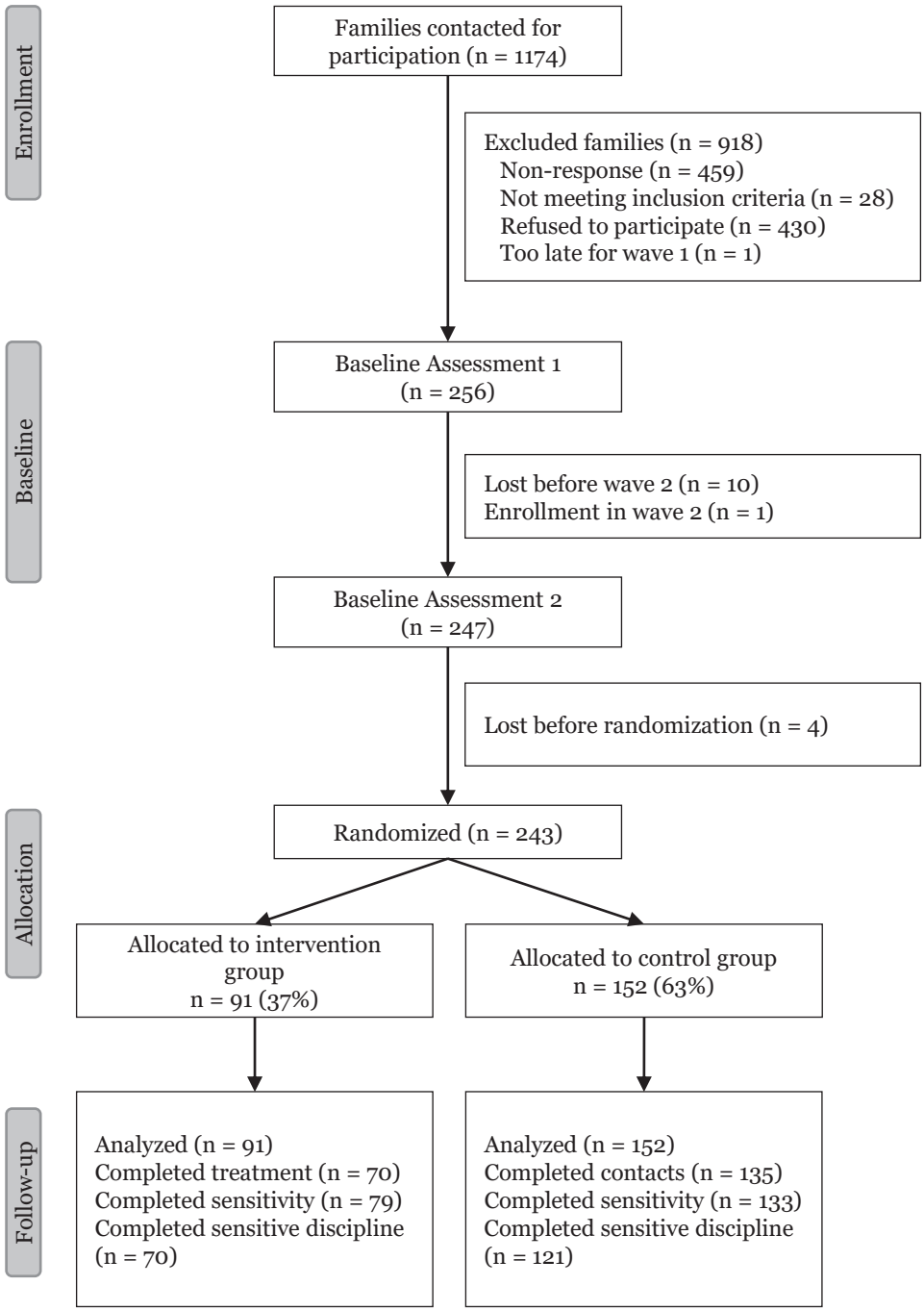
Table 5.1

Characteristics of the sample, and separately for intervention & control groups

	Total	Inter-vention	Control
	N = 243	N = 91	N = 152
Twin characteristics			
Age at T1 in years M (SD)	7.92 (0.66)	7.94 (0.66)	7.92 (0.67)
Sex (% boys)	48.6	49.5	48.0
Country of birth (% Netherlands)	99.2	100	98.7
Zygosity (% MZ)	55.1	50.5	57.9
Family characteristics			
Primary parent (%)			
Biological mother	90.5	87.9	92.1
Adoptive Mother	0.8	1.1	0.7
Biological father	8.6	11.0	7.2
Age primary parent M (SD)	40.48 (4.66)	40.77 (4.78)	40.32 (4.60)
Age second parent M (SD)	42.37 (5.33)	42.56 (5.33)	42.25 (5.35)
Country of birth (% Netherlands)			
Primary parent	97.5	96.7	98.0
Second parent	95.5	95.6	96.0
Educational level primary parent			
Lower and Intermediate vocational	34.3	35.2	33.8
Higher vocational, university bachelor	41.7	39.6	43.0
Post-higher vocational, university master	24.0	25.3	23.2
Number of other children in the family M (SD)	1.40 (0.68)	1.33 (0.61)	1.44 (0.73)
Number of older siblings M (SD)	0.81 (0.87)	0.86 (0.82)	0.77 (0.89)
Number of younger siblings M (SD)	0.22 (0.50)	0.18 (0.41)	0.25 (0.54)
Primary parents' marital status (%)			
Married or registered partnership	74.0	67.0	78.3
Cohabiting	19.8	26.4	15.8
Single parent	6.2	6.6	5.9
Parental temperament M (SD)	4.31 (0.62)	4.26 (0.59)	4.33 (0.64)
Parental psychopathological symptoms T1	0.20 (0.24)	0.19 (0.19)	0.21 (0.27)
Parental psychopathological symptoms T2	0.30 (0.42)	0.26 (0.30)	0.32 (0.47)
Parental psychopathological symptoms T3	0.26 (0.28)	0.26 (0.26)	0.26 (0.29)

Figure 5.1

Flow Chart Recruitment, Exclusion, and Randomization of the Study Sample



In order to fit the intervention to parents of older children, some slight adaptations were made. For example, in the second visit the discipline situation "clearing up" was replaced by the more age adequate Family Interaction Task (FIT, Allen et al., 2005). Parents and children were asked to choose one topic of a list with issues about which they might argue (e.g., bedtimes, clothing or media use), discuss their disagreements and to come up with ideas on how to resolve these issues. When discussing the recordings of this task, we advised parents to involve their children in setting up house rules and gave some tips on how to do this. Furthermore, when reflecting on the sensitive time-out in Session 4, we added information on how to handle the situation when the child instead of the parent initiates a time-out by walking away. Other changes involved age-adequate play materials (e.g., a board game). As is common in an RCT testing the effectiveness of the VIPP-SD, the five interveners were extensively trained following a standard training of 4 full days and an additional half day to be trained in the adaptations of the twin intervention. Together with the trainer, they practiced the program in a family with twins. Moreover, interveners kept logs about adherence to the protocol (which they did not deviate from) and attended regular intervision sessions (every second or third week) to support interveners in the implementation of the intervention according to protocol and to monitor the fidelity and quality of the intervention.

Control Condition. Families in the control condition received five phone calls parallel to the intervention sessions to assure that they had the same number of contacts. Following a standard protocol, families were asked general and specific questions about the development of their twins using a semistructured interview. We provided no information or advice about parenting. When parents asked for advice on parenting issues, they were referred to online information about parenting. If they asked for advice on parenting of twins specifically, they were referred to the Dutch organization for parents with multiples (Nederlandse Vereniging voor Ouders van Meerlingen; NVOM).

Measures

Observed Parental Sensitivity. We observed sensitivity of the primary parent with each twin separately, using a computerized version of the Etch-A-Sketch task

which is unknown to contemporary Dutch families (Cents et al., 2014). The order of the performance of the task ("oldest" child first or "youngest" child first) was randomly chosen. The parent-child dyad had to imitate three printed example drawings (with increasing difficulty) on a computer screen (Euser et al., 2021). The interaction between the parent and child was filmed and the screen with the drawings was recorded. We coded sensitivity using the revised Erickson (Egeland, 1990) 7-point rating scales in concordance with the Euser et al. (2021) procedure. The two scales comprise supportive presence (1 = parent completely fails to be supportive to the child, 7 = parent skillfully provides support throughout the session) and intrusiveness (1 = parent allows the child sufficient time to explore and to attempt to solve tasks on her/his own, 7 = parent is highly intrusive; her/his agenda clearly has precedence over the child's wishes). Thirteen coders, trained by two expert coders (MBK and SE), coded the videos. Intercoder reliability (intraclass correlation coefficient; ICC) with the expert coder and among coders was adequate at T1, T2, and T3 (ICC = .71–.78, see Table S5.1). We reversed intrusiveness scores; therefore, higher parental sensitivity was indicated by higher scores on both scales. The correlation between the two scales was high at each time point (r = .53–.61, see Table S5.1), therefore (and in accordance with the Euser et al. (2021) study), the scores of both scales were averaged.

Observed Parental Sensitive Discipline. We observed sensitive discipline of the primary parent using an adapted version of the Do-Don't task (Kochanska, 1995; Van Der Mark et al., 2002; Yagmur et al., 2014). The order of children was randomly chosen. The parent received a laptop with written instructions and the child received materials needed for the task. First, parents had to explain the task to their child. Then, parents were asked to watch a video during the task of the child and to ensure that the child would not watch the video. Thus, the parents had to set limits in two different ways: Ensuring that the child would engage in the not so attractive "do" task and preventing the child from engaging in the attractive "don't" task. At T1, the "do" task consisted of color sorting loombands in a box, at T2 children had to color sort perlerbeads and at T3 children had to do a writing task. At all three timepoints, the parents watched an attractive child movie/series (*Buurman & Buurman* at T1 and T2 and *the Minions* at T3). The task lasted for 8

min and was filmed.

We used the positive discipline scale to measure sensitive discipline in accordance with the Euser et al. (2021) study. Positive discipline was rated on an adapted version of the revised Erickson 7-point rating scale (Egeland, 1990) for supportive presence (1 = parent completely fails to provide positive discipline, 7 = parent skillfully provides positive discipline throughout the session). In addition, parental physical interference and laxness were rated but not used in the analyses because of highly skewed score distributions (and in accordance with Euser et al., 2021). Seventeen coders, trained by two expert coders (JR and SE), coded the videos. Intercoder reliability (intraclass correlation coefficient; ICC) with the expert coder and among coders was adequate at T1, T2, and T3 (ICC = .73–.88, see Table S5.1).

Parental Attitudes Toward Sensitivity and Sensitive Discipline. At T1, T2, and T3, parents filled in the Dutch version of the Parenting Attitudes Questionnaire (PAQ) consisting of 20 items (Van Zeijl et al., 2006) which are scored on a scale from 0 to 100. For each item, parents indicated their position on a 10-cm line ranging from totally disagree to totally agree. We conducted a principal component analysis with a varimax rotation. Eight items loaded onto the sensitivity factor and ten items loaded onto the discipline factor, two items did not load on any of the factors. Therefore, we computed two scale scores, one for sensitivity based on eight items and one for discipline based on ten items. Factor loadings were not equal across measurements, in which case it is recommended to report McDonalds' omega instead of α (Viladrich et al., 2017). Over the three waves, omega ranged from .60 to .62 for sensitivity and from .59 to .67 for sensitive discipline. Example items for sensitivity were "In my opinion, I should praise my child at least once every day" and for discipline were "My child must learn that I will get angry when he/she does not listen to me" (reversed).

Temperamental Reactivity of the Primary Parent. We used the Dutch 15-item orienting sensitivity scale from the adult temperament questionnaire (short form) to measure susceptibility to the environment in accordance with Euser et al. (2021). The sensitivity scale is comprised of three subscales: Neutral perceptual sensitivity, affective perceptual sensitivity, and associative sensitivity (Evans & Rothbart, 2007). Items were measured on a 7-point Likert scale (1 =

extremely untrue to 7 = extremely true). Reliability for the self-report and partner report was satisfactory (Cronbach's α = .72 for the self-report and α = .76 for the partner-report). Self-reported and partner-reported temperamental reactivity correlated significantly (r = .33), therefore the mean of both scores was used. Data was missing completely at random (MCAR) for 33 families (12.8%). We imputed the data using the expected means (EM) function in SPSS with age and gender of the primary parent, age and gender of the twin, parental educational level, and parental psychopathology symptoms as predictors. The imputed data was normally distributed.

Psychopathological Symptoms of the Primary Parent. We administered four subscales of the Dutch version of the Brief Symptom Inventory (BSI) to control for possible psychopathological differences between parents in the intervention group and control group (Derogastis, 1993): Anxiety (six items), depression (six items), hostility (five items), and interpersonal sensitivity (four items). A mean score of all 21 items was computed (higher scores indicating more symptoms). Cronbach's α was good at all time points (α = .87 at T1, α = .93 at T2, and α = .86 at T3). There were 12% of data missing at T1, 6.9% at T2, and 13.9% at T3. Missing data occurred completely at random according to the MCAR test, and we imputed missing data using the EM function in SPSS with age and sex of the primary parent and their children, parental education level and the BSI scores as predictors. The imputed scale scores were square-root transformed.

Data Analyses

We used the same statistical procedures as in Euser et al. (2021), conducting intent-to-treat analyses. We applied multilevel analyses with full maximum likelihood estimation due to the nature of the data (longitudinal and nested). We had three levels (time of assessment, child, family) in the analysis of observed sensitivity and sensitive discipline and two levels (time of assessment, parent) in the analysis of attitudes toward sensitivity and sensitive discipline. We computed intraclass correlation coefficients to investigate whether the levels explained a significant proportion of the variance. Four separate models were examined, one for each outcome variable (parental sensitivity, parental sensitive discipline, attitudes

toward sensitivity, attitudes toward sensitive discipline). In a first step, we computed the ICCs by fitting an intercept only model with three levels (or two levels) and an unrestricted within-subject (co)variance structure. Then we fitted a growth model. Time (coded as 1–3) and Time^2 (coded as 1, 4, and 9) were added as fixed and random effects. Time^2 was added because the intervention occurred between T2 and T3, so we could test for a quadratic effect. Condition (coded as 0 for the intervention group and 1 for the control group) and Condition × Time^2 were included as fixed effects to test the intervention effect. As a last step, Time^2 × Condition × Parental temperamental reactivity was added as fixed effect to test whether parents were differentially susceptible to the intervention program. We standardized all predictor variables before inclusion. In robustness analyses, we conducted the same analyses as mentioned above in participants who received all five intervention sessions/ control calls, and in participants with complete data on the outcome variables at all assessments. Next, we examined the effects of the VIPP-SD on parental sensitivity and sensitive discipline in depth using an exploratory approach. For sensitivity, we distinguished between the two scales (supportive presence and intrusiveness) to assess whether the intervention had differential effects on different dimensions of parenting. Second, we investigated in more detail whether children within families might trigger different parenting intervention effects. Therefore, we reran the analyses first including the sensitivity or sensitive discipline scores at the first measurement as covariate and then including zygosity as a covariate to test whether intervention effectiveness was dependent on the baseline level of parenting. The a-priori power with an α of .05 and an expected mean difference of 0.10 was good (.80) considering a sample size of 243 families.

Transparency and Openness

We report all data inclusions or exclusions as well as all measures. The analysis code and research materials are available upon request. Data were analyzed using SPSS (Version 27). The trial, the study design, and the analysis were pre-registered. For the analyses exploring whether baseline differences and temperamental reactivity moderated the effects, we decided to use continuous instead of dichotomized variables. We also decided to include parental attitudes in

the current report. These adaptations were reported in a time-locked addendum to the original pre-registration.

Results

Preliminary Analyses

Randomization was checked by comparing the intervention group and control group on all background variables. The groups did not differ significantly on any of the variables. Therefore, we did not include any background variables as covariate. Table 5.1 shows all variables, their means or frequency for the total group as well as for the intervention and control group separately. Treatment integrity was high, 83% of the families participated in all intervention sessions or dummy contacts. Table 5.2 shows descriptive statistics of the outcome variables. The correlations within twin and across time were significant: Within twin correlations ranged from $r = .42$ to $r = .57$ for parental sensitivity and from $r = .38$ to $r = .66$ for sensitive discipline, across wave correlations ranged from $r = .41$ to $r = .46$ for parental sensitivity and from $r = .15$ to $r = .37$ for sensitive discipline.

Primary Intervention Effects

Sensitivity. The ICC in the intercept only model with three levels revealed that 42% of the variance in parental sensitivity could be attributed to the family, meaning that sensitivity scores within a family were more similar to each other than sensitivity scores between families. Four percent could be attributed to the child level, meaning that the sensitivity scores at T1, T2, and T3 within a child were not much more similar compared to scores from the twin sibling. In subsequent analyses, we therefore included only family and time as levels. Table 5.3 shows the statistics for testing a main effect of the intervention and a moderating effect of temperamental reactivity on parental sensitivity. Time and quadratic time were significant (time: $\beta = 0.57$, $p = .002$ and quadratic time: $\beta = -0.53$, $p = .007$) implying change in parental sensitivity levels over time. The interaction between condition and quadratic time was not significant suggesting that the intervention did not affect parental sensitivity. A main effect of temperamental reactivity was found ($\beta = 0.31$, $p = .002$) with parents scoring higher on temperamental reactivity

having higher sensitivity scores. The interaction between condition, reactivity, and quadratic time was not significant, implying that parents' temperamental reactivity did not moderate intervention effectiveness. The robustness analysis with cases with complete treatment and the robustness analysis with cases with complete assessment showed similar results (see Table S5.2).

Sensitive Discipline. The ICC in the intercept only model with three levels revealed that 34% of the variance in parental sensitivity could be attributed to the family, whereas 0% could be attributed to the child level. In subsequent analyses, we therefore included only family and time as levels. Table 5.3 shows the statistics for the main effect of the intervention and the moderating effect of temperamental reactivity on intervention effects on parental discipline. Time and quadratic time were significant (time: $\beta = -1.25$, $p < .001$ and quadratic time: $\beta = 1.05$, $p < .001$). All other effect were not significant, suggesting that the intervention did not influence parental sensitive discipline and the intervention effect was not moderated by parents' temperamental reactivity. The robustness analysis with cases with complete treatment and the robustness analysis with cases with complete assessment showed similar results (see Table S5.2). It should be noted, however, that a null finding in one replication study does not mean that the finding of the original study is disconfirmed or falsified. The original result is a point estimate with a boundary around it which is called the "prediction interval". This is the interval in which the point estimate of a next study will be positioned with a high chance (Margoni & Shepperd, 2020). For the original study with an estimated main effect size of $r = .15$ for sensitive discipline (Euser et al., 2021) the 95% prediction interval is PI [−0.03, 0.33] (see https://replication.shinyapps.io/correlation/) based on the current replication sample size of $N = 243$. If the effect size of the replication study differs from the effect size of the original study only due to sampling error, there is a 95% chance that the replication result will fall in this interval (Spence & Stanley, 2016). The effect size of our null finding for parental sensitive discipline is within the prediction interval of the original study, thus probably due to sampling error instead of systematic influences or differences between the two studies, for example age differences.

Secondary Intervention Effects

Attitudes Toward Sensitivity. The ICC in the intercept only model with two levels revealed that 54% of the variance in attitudes toward sensitivity could be attributed to the parent, meaning that the scores within a parent were more similar to each other than scores between parents. Table 5.4 shows the statistics for testing a main effect of the intervention and a moderating effect of temperamental reactivity on attitudes toward sensitivity. Quadratic time was significant ($\beta = 5.04$, $p = .021$). The interaction between condition and quadratic time was significant ($\beta = 3.44$, $p < .001$, Cohen's d = 0.26) indicating that the intervention increased positive attitudes toward sensitivity significantly more in the intervention group than in the control group (Figure 5.2). The interaction between condition and temperamental reactivity was not significant. The three-way interaction between condition, quadratic time, and temperamental reactivity was not significant either. The robustness analyses with cases with complete treatment and the robustness analysis with cases with complete assessment showed similar results (see Table S5.2).

Attitudes Toward Sensitive Discipline. The ICC in the intercept only model with two levels revealed that 58% of the variance in parental sensitivity could be attributed to the parent, meaning that the scores within a parent were more similar to each other than scores between parents. No significant effects were found (see Table 5.4). The robustness analysis with cases with complete treatment and the robustness analysis with cases with complete assessment showed similar results (see Table S5.2).

Exploratory Analyses

For sensitivity, we kept the two scales for observed parental sensitivity (supportive presence and nonintrusiveness) apart to assess whether the intervention had differential effects on different dimensions of sensitive parenting. We also reran the analyses with the mean sensitivity score (or mean sensitive discipline score) on the first measurement (T1) as a covariate. Lastly, we reran the analyses with zygosity of the children as covariate. All results did not substantially deviate from the results of the main analyses (see Table S5.3).

Table 5.2

Means and standard deviations of parental sensitivity, sensitive discipline and parental attitudes towards sensitive discipline at baseline 1 (T1), baseline 2 (T2), and post-test assessment (T3) for the intervention and control group

	T1		T2		T3	
	Child 1	**Child 2**	**Child 1**	**Child 2**	**Child 1**	**Child 2**
Observed parental sensitivity						
Intervention	3.73 (1.24)	3.51 (1.32)	4.00 (1.28)	3.86 (1.32)	3.75 (1.28)	3.68 (1.15)
Control	3.87 (1.25)	3.71 (1.29)	3.94 (1.27)	3.92 (1.22)	3.86 (1.33)	3.95 (1.29)
Observed Parental sensitive discipline						
Intervention	4.30 (1.72)	4.19 (1.62)	3.87 (1.58)	3.67 (1.62)	3.61 (1.28)	3.76 (1.31)
Control	4.11 (1.38)	4.29 (1.40)	3.58 (1.52)	3.54 (1.44)	3.52 (1.30)	3.84 (1.42)
Parental attitudes towards sensitivity						
Intervention	65.14 (14.29)		65.75 (13.28)		74.19 (12.95)	
Control	66.38 (13.12)		66.87 (14.36)		69.12 (11.88)	
Parental attitudes towards sensitive discipline						
Intervention	60.23 (12.81)		59.54 (12.41)		61.18 (12.21)	
Control	60.41 (11.96)		58.83 (13.82)		58.88 (11.76)	

Table 5.3

Multilevel model statistics testing the intervention effect and moderator effect on parental sensitivity and sensitive discipline

	Parental Sensitivity				Parental Sensitive Discipline			
	Est	**SE**	**p**	**95% CIs**	**Est**	**SE**	**p**	**95% CIs**
Intercept	3.80	0.10	.000	3.61 – 3.98	3.82	0.80	.000	3.66 – 3.98
Time	0.57	0.19	.002	0.20 – 0.94	-1.25	0.26	.000	-1.76 – -0.73
Time^2	-0.53	0.19	.007	-0.91 – -0.15	1.05	0.26	.000	0.54 – 1.57
Condition	0.07	0.12	.585	-0.17 – 0.30	0.08	0.13	.521	-0.17 – 0.34
Condition*Time2	0.01	0.06	.921	-0.11 – 0.12	-0.05	0.08	.479	-0.20 – 0.10
Reactivity	0.31	0.10	.002	0.12 – 0.51	0.06	0.08	.405	-0.09 – 0.22
Condition*Reactivity	-0.23	0.12	.058	-0.48 – 0.01	0.20	0.13	.131	-0.06 – 0.46
Reactivity*Time^2	-0.16	0.05	.735	-0.11 – 0.08	-0.07	0.05	.132	-0.16 – 0.02
Condition*Reactivity*Time^2	0.04	0.06	.520	-0.08 – 0.15	-0.03	0.08	.712	-0.18 – 0.12

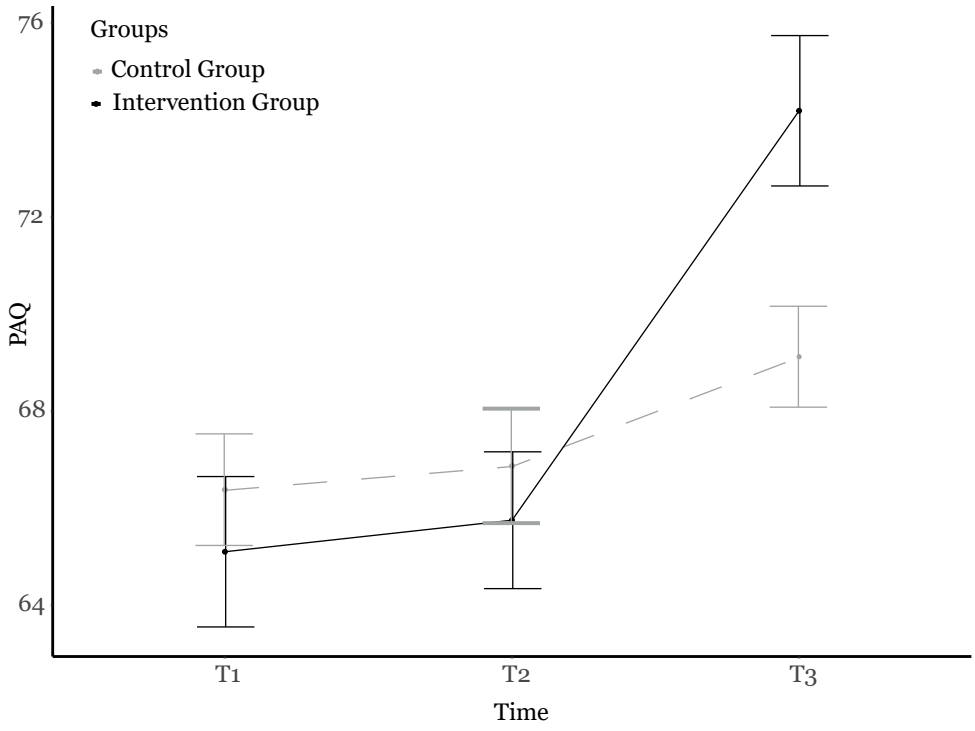
Table 5.4

Multilevel model statistics testing the intervention effect and moderator effect on parental sensitivity and sensitive discipline

	Parental Sensitivity				Parental Sensitive Discipline			
	Est	**SE**	**p**	**95% CIs**	**Est**	**SE**	**p**	**95% CIs**
Intercept	67.31	0.84	.000	65.65 – 68.96	59.50	0.79	.000	57.95 – 61.06
Time	-4.04	2.20	.067	-8.36 – 0.29	-3.44	2.04	.093	-7.47 – -0.58
Time^2	5.04	2.17	.021	0.78 – 9.30	2.79	2.03	.169	-1.20 – 6.78
Condition	1.45	1.38	.293	-1.26 – 4.16	0.84	1.30	.516	-1.71 – 3.39
Condition*Time2	3.44	0.85	.000	1.73 – 5.15	1.07	0.80	.182	-0.51 – 2.65
Reactivity	0.55	0.83	.504	-1.08 – 2.18	0.24	0.78	.754	-1.29 – 1.78
Condition*Reactivity	1.40	1.41	.321	-1.37 – 4.18	1.18	1.33	.374	-1.42 – 3.79
Reactivity*Time^2	0.78	0.53	.147	-0.28 – 1.83	-0.68	0.49	.171	-0.65 – 0.29
Condition*Reactivity*Time^2	-0.78	0.87	.374	-2.53 – 0.97	0.87	0.81	.285	-0.74 – 2.48

Figure 5.2

Multilevel model statistics testing the intervention effect and moderator effect on parental sensitivity and sensitive discipline



Discussion

In this replication study, we investigated the effects of the VIPP-SD on parental sensitivity and sensitive discipline in parents of twin children in middle childhood. We did not find a significant effect of the intervention on sensitivity or sensitive discipline. We also investigated whether the VIPP-SD had an effect on parents' attitudes toward sensitivity and sensitive discipline. The VIPP-SD had a significant effect on parents' attitudes toward sensitivity but not on parent's attitudes toward sensitive discipline. We tested whether parents were differentially sensitive and differed in their disciplining behavior with each of their twin children, but this was not the case. Based on the idea of differential susceptibility, we further examined whether parents were differentially susceptible to the intervention effects depending on their temperamental reactivity. We did not find support for the differential susceptibility hypothesis. Contrary to our expectations, the VIPP-SD did not influence parental sensitivity as assessed in the present study. Although this is in line with the findings of the Euser et al. (2021) study, it is not what we had expected based on meta-analytic findings on the effectiveness of the VIPP-SD program (Juffer et al., 2017; Van IJzendoorn et al., 2023). It might be that the intervention indeed failed to have an effect on sensitivity. Parents of school-aged children have many years of experience in raising their children and parenting behavior that developed over such a long time might be harder to change than interactive behavior of parents with younger children. Originally, the VIPP-SD was designed for children up to 3 years. In the preschooler study with children between 5 and 7 years, the effect on sensitivity also did not emerge anymore (Euser et al., 2021). The effectiveness of the VIPP-SD might therefore be limited to infancy and early childhood, at least in upper- and middle-class families with twins. The IPD meta-analysis of Incredible Years interventions by Gardner et al. (2019) showed similar child effects of interventions between the ages of 2 and 12. Although children might be equally responsive to interventions at several stages of their childhood, the same might not be true for their parents. This might mean that parents of school-aged twins may need more than five sessions for behavioral change to happen or that the intervention is indeed not effective for parents of school-aged children. In the latter case, if the null finding is replicated in follow-up studies and other RCTs, we need to be cautious to not put extra burden

on similar participants by continuing to test a (possibly) ineffective intervention. Another explanation is the nature of the task in which parental sensitivity was observed. Only one of the 12 studies in the meta analysis about the effectiveness of the VIPP-SD (Juffer et al., 2017) used a challenging problem-solving puzzle, which can be considered comparable to the etch-a-sketch task we used. Van Zeijl et al. (2006), who used this challenging task did find significant effects on the attitude of mothers toward sensitivity but not on maternal sensitivity itself. In our study we used a challenging task for the parents and the children which was impossible to master. They had to draw three pictures including circles in a program where no circles could be drawn, making it a frustrating activity. Parents' sensitivity may have increased due to the intervention but might still have been insufficient for the setting in which sensitivity was assessed at the posttest of the present study. Future research might incorporate other tasks, such as a discussion with distinct emotional valence. Although Euser et al. (2021) found an effect of the VIPP-SD on sensitive discipline, we did not. As with parental sensitivity, the age of the children might be an explanation. The VIPP-SD has been found to be effective in promoting sensitive discipline in samples with young children and also in the preschooler sample, but not later, so the upper age limit of the VIPP-SD might be reached after 7 years of age. The discrepancy with the preschooler study (Euser et al., 2021) might also be explained by difference in the task we used to observe parental sensitive discipline. In our study, we used the combined Do-Don't task to assess sensitive discipline. In the preschooler study, there was only a Don't Touch task, where a child was not allowed to touch a range of toys. The parent was free to focus on his or her child and had only one task, but in our study the parents had to (a) make sure that the child stayed engaged in the "Do" task, (b) make sure the child did not (secretly) watch the video ("Don't"), and (c) watch the video themselves. The demands of the task might have been too complicated to utilize their newly learned parenting skills. In line with Van Zeijl et al. (2006), we found a significant effect of the VIPP-SD on the attitudes of parents toward sensitivity which indicates that the intervention did influence the attitudes of parents in a positive direction. Promoting changes in parental cognitions about sensitivity is an important goal of VIPP-SD and parent coaching in general. According to Bandura's (1977) theory, cognitions, (such as "sensitive interactions are

desirable and can be reached") are theorized to be part of mechanisms of change in which performance feedback and modeling through video-feedback lead to reciprocal changes in efficacy beliefs and performance of target behavior in the longer term (Schuengel & Oosterman, 2019). Changed attitudes may be the first step toward a behavioral change (Bakermans-Kranenburg & Oosterman, 2021). A sleeper effect on sensitive parenting behavior may be observed at later timepoints. Future research should investigate if and how much time is needed for parental changed attitudes to be translated into changed behavior.

We did not find support for the differential susceptibility hypothesis regarding the intervention effects. Parents who were more temperamentally reactive did not benefit more from the VIPP-SD than parents with lower reactivity, convergent with the preschooler twin study (Euser et al., 2021). In general, most differential susceptibility studies focus on children as subjects (Ellis et al., 2011) and temperament has more often been used as a differential susceptibility marker in children, but not in adults (Belsky & van IJzendoorn, 2017). For adults, susceptibility markers that have been examined were mostly dopamine and serotonin polymorphisms (van IJzendoorn & Bakermans-Kranenburg, 2015). This suggests that reactive temperament is a suitable phenotypical marker of susceptibility in childhood, but maybe not in adulthood. Future studies might aim to test differential susceptibility to parenting intervention efforts using neurobiological and genetic markers to assess if these are more suitable susceptibility marker than reactive temperament in adults. Parental sensitivity and sensitive discipline were quite similar for both children within a family, and this was also the case in the group of families with DZ twins. Our findings are congruent with the preschooler twin study (Euser et al., 2021) in which parental sensitivity and sensitive discipline were similar for both children in a family too. Both studies involved same-sex twins, thus our finding cannot be extended to twins of different sex, and to siblings of different ages or sex, which poses an important set of questions for future research.

Our study has strengths and limitations. We pre-registered our analyses to increase the reproducibility and transparency of our results. Furthermore, the study design was a randomized controlled trial with a large sample and two pretests. In parenting research, bidirectional effects have been emphasized in which parents'

behavior might influence the child but at the same time the child might affect the behavior of the parent as well. A RCT is a particular strong method to test causality as well as the direction of an association between parenting and child outcomes, although the quality of RCTs as well as their generalizability vary across interventions, for example when self-reports are used (Bakermans-Kranenburg et al., 2003; Hamaker et al., 2020). We assessed parents' attitudes using a questionnaire with modest internal consistency. Future studies may improve on measures for assessing parents' attitudes and assessing sensitivity and discipline in a variety of more and less challenging parent–child interaction settings. Furthermore, we assessed parenting using observational measures of sensitivity and sensitive discipline but in specific, maybe too demanding settings. Our measure of parental temperament included a partner-report questionnaire. Doing so we increased reliability, but we did not include a potentially even more valid measure, such as a polygenic score as susceptibility marker. Such a polygenic susceptibility score has been derived from a genome-wide association study (GWAS) by Keers et al. (2016) using differences in development of internalizing problems between monozygotic (MZ) twins. This polygenic score has been applied with some success by Lemery-Chalfant et al. (2018) in their randomized intervention study. Polygenic susceptibility scores go beyond the candidate genes approach that often only explains a small amount of variance. It uses a combination of thousands of single-nucleotide polymorphisms (SNPs) derived from GWAS data characteristic of MZ twins who respond differently to the environment, thus suggesting differential susceptibility (Belsky & van IJzendoorn, 2017). Last, to assess whether the intervention influences parenting on the long term, posttest assessments at a later moment in time after the intervention are needed to detect potential sleeper effects. The positive effect of the VIPP-SD on parental attitudes toward sensitivity may only after some time translate into more sensitive behavior.

Conclusion

In conclusion, we did not find a significant effect of the VIPP-SD on parental sensitivity or sensitive discipline in parents of schoolaged twins. One possible explanation is that we have reached the upper age limit for this intervention,

originally developed for infants and preschoolers, to be effective. Alternatively, the VIPP-SD might not have increased sensitivity in such a way that highly demanding interaction settings with school-age children become manageable for parents in the short run. However, we did find that the VIPP-SD improved parents' attitudes toward sensitivity and parents might need more time to translate their changed attitudes into parenting behavior.

Acknowledgements

We thank the participating families for their enthusiastic involvement in the Leiden Consortium on Individual Development (L-CID). We are also grateful to the data-collection and data-processing team, including all current and former students, research assistants, PhD students and post-doctoral researchers for their dedicated and invaluable contributions. Marinus van IJzendoorn, Eveline Crone and Marian Bakermans-Kranenburg designed the L-CID experimental cohort-sequential twin study "Samen Uniek" as part of the Consortium on Individual Development (CID; Gravitation Grant 2013-2023 awarded by the Dutch Ministry of Education, Culture, & Science, and the Netherlands Organization for Scientific Research, NWO Grant Number 024.001.003). We thank Saskia Euser for training observers to code sensitive caregiving, discipline, and SBSK.

Supplementary Materials

All supplementary materials are published and can be retrieved from:
http://tinyurl.com/Chapter5Supplement

This chapter is published as:

Runze, J., Pappa, I., Van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2022). Conduct Problems and Hair Cortisol Concentrations Decrease in School-Aged Children after VIPP-SD: A Randomized Controlled Trial in Two Twin Cohorts. *International Journal of Environmental Research and Public Health, 19*(22), 15026. https://doi.org/10.3390/ijerph192215026

CONDUCT PROBLEMS AND HAIR CORTISOL CONCENTRATIONS DECREASE IN SCHOOL-AGED CHILDREN AFTER VIPP-SD: A RCT IN TWO TWIN COHORTS

CHAPTER SIX

Abstract

The Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD) is effective in increasing parental sensitivity and sensitive discipline, and aims to decrease child behavior problems. Changes in quality of parenting may be accompanied by effects on child stress levels. However, studies of VIPP-SD effects on child behavior problems have shown mixed results and there are no studies to date of the effect of the intervention on children's stress levels, as measured by hair cortisol concentration (HCC). Furthermore, differences in intervention effectiveness may be explained by differential susceptibility factors. We hypothesized that the effects of the VIPP-SD on child behavior problems might be moderated by currently available child polygenic scores of differential susceptibility (PGS-DS). In the current pre-registered trial, we randomly assigned 40% of n = 445 families with school-aged twin children to the intervention group. The VIPP-SD was successful in decreasing both children's conduct problems and HCC. Effects were not moderated by available child PGS-DS. We conclude that a brief, home-based video-feedback parenting intervention can decrease child behavior problems and affect the child's stress-related neuroendocrine system as assessed with hair cortisol. In future studies, more specific PGS-DS for externalizing behaviors should be used as well as parental PGS-DS.

Meta-analytic evidence has shown that the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD) is effective in increasing parental sensitivity, parental sensitive discipline, and child attachment security (Juffer et al., 2017; van IJzendoorn et al., 2023). Parenting and attachment are important factors in the development and continuity of children's stress regulation and externalizing behavior problems (Groh et al., 2017; Gunnar, 2017; Miner & Clarke-Stewart, 2008). High-quality parenting and secure attachments have been found to predict lower levels of externalizing behaviors (Cooke et al., 2022) and to buffer the stress reactivity of the hypothalamic–pituitary–adrenocortical (HPA) axis (Gunnar, 2017; Hostinar et al., 2014; Simmons et al., 2019). In the current randomized controlled trial (RCT) in two cohorts of families with twins, we therefore examined the effects of the VIPP-SD on children's conduct problems and hair cortisol concentrations (HCC).

During the VIPP-SD, parents learn tools such as distraction and inductive discipline, in order to respond adequately to difficult child behavior, which helps decreasing problem behavior in the long term (Juffer et al., 2017). Moreover, parents learn to see even subtle signals of proximity-seeking behavior in their child and to react promptly and adequately, which should help decreasing child stress levels (Juffer et al., 2017). Although the effects of the VIPP-SD on parental sensitivity and secure attachment have been supported in several randomized trials and confirmed in a meta-analysis of 25 randomized controlled VIPP-SD studies, the evidence for the effects of VIPP-SD on children's externalizing behavior remains inconclusive (van IJzendoorn et al., 2023). The most recent and largest pragmatic randomized trial conducted within the context of the British National Health Service (NHS) including more than 300 families with toddlers at risk for externalizing problems showed a significant effect of VIPP-SD on conduct problems (O'Farrelly et al., 2021). Similarly, some previous research showed the effects of parenting interventions on children's stress regulation but meta-analytic results were equivocal. The Attachment and Biobehavioral Catch-up (ABC) intervention with 4- to 6-year-old children involved with Child Protective Services (CPS) resulted in more typical salivary diurnal cortisol levels in the intervention group (Bernard et al., 2015). Also, in 1- to 3-year-old children with high externalizing behavior, diurnal cortisol production as measured

in saliva samples decreased after the VIPP-SD, specifically in carriers of the DRD4 7-allele (Bakermans-Kranenburg & van IJzendoorn, 2008). However, in a recent meta-analysis of 19 parenting intervention studies, no overall significant effect of parenting interventions on child salivary cortisol levels was found (Martins et al., 2020). Further work is needed to provide more robust evidence for the effectiveness of attachment-based parenting interventions on child externalizing behavior and stress regulation.

An avenue not yet explored in parenting intervention studies on hormonal stress regulation is the use of hair cortisol concentrations (HCC, Stalder & Kirschbaum, 2012). In developmental research, cortisol has mostly been measured using saliva, but salivary cortisol only provides a snapshot of cortisol levels on a particular day (Kirschbaum & Hellhammer, 1994). However, cortisol levels may vary strongly from day to day as they are influenced by many factors, such as food intake and activity (Strahler et al., 2017). Furthermore, only chronic stress indicated by long-term deviations in cortisol levels as assessed with HCC across several months may be problematic (Chrousos & Kino, 2007). To our knowledge, only one small RCT including 25 parents and their children investigated the effect of compassion training in parents on children's HCC and found decreased HCC in the intervention group (Poehlmann-Tynan et al., 2020).

Another topic to explore more thoroughly in parenting interventions is the moderating role of children's genetic differential susceptibility to parenting (Belsky et al., 2007; Ellis et al., 2011). Highly susceptible children are supposed to show more psychological problems when exposed to adverse environments, but also to benefit more from (experimentally) enhanced caregiving environments (Bakermans-Kranenburg & van IJzendoorn, 2015). A promising approach to genetic differential susceptibility is the use of polygenic scores (PGSs) derived from genome-wide association studies (GWAS) to compute markers of differential susceptibility (Belsky & van IJzendoorn, 2017; Keers et al., 2016). Keers and colleagues (2016) were the first to develop a genome-wide association-based PGS for differential susceptibility to anxiogenic environments. Based on within-pair variability in emotional problems in monozygotic twins, they identified SNPs of environmental susceptibility and combined these in a PGS. This PGS was used to test the differential effects of

treatment (Keers et al., 2016) and a Family Check-Up intervention on children's internalizing symptoms (Lemery-Chalfant et al., 2018). In the absence of a specific PGS for susceptibility to externalizing-inducing environments, we decided to use the PGS for anxiogenic environments as a possible moderator of VIPP-SD effects as externalizing and internalizing problems usually show substantial correlations. In addition, for exploratory purposes, we computed a PGS based on a GWAS of susceptibility to environmental stress and adversity in adults (SESA) (Nagel et al., 2020) that might also be a proxy for differential susceptibility to parenting.

In sum, we implemented the VIPP-SD intervention in a randomized controlled trial on two twin samples (Leiden Consortium on Individual Development (L-CID), total n = 890) measuring children's conduct problems and HCC before and after the intervention. We expected that, after the intervention, children in the intervention group would exhibit fewer conduct problems compared to children in the control group (H1). Furthermore, we expected that, after the intervention, children's HCC would be significantly lower in the intervention group as compared to the control group (H2). Lastly, we hypothesized that the intervention effects on HCC and conduct problems would be moderated by putative child polygenic scores for differential susceptibility (H3), with stronger intervention effects for children with higher polygenic scores.

Materials and Methods

Participants

Participants were part of the L-CID project, a longitudinal intervention study with families with twins belonging to one of two cohorts: an early childhood cohort (n = 237 families) and a middle childhood cohort (n = 256 families) (Figure 6.1). At the first measurement, the children of the early childhood cohort were on average 3.76 years old (*SD* = 0.57) and the children of the middle childhood cohort were on average 7.92 years old (*SD* = 0.66). Sixty percent of the children were monozygotic twins and forty-five percent were male. Descriptive statistics for both cohorts and the full sample can be found in Table 6.1.

Figure 6.1

Flow chart of the randomized controlled trial with VIPP-SD in both twin cohorts. Number of participants reflects the number of families (including two twin children).

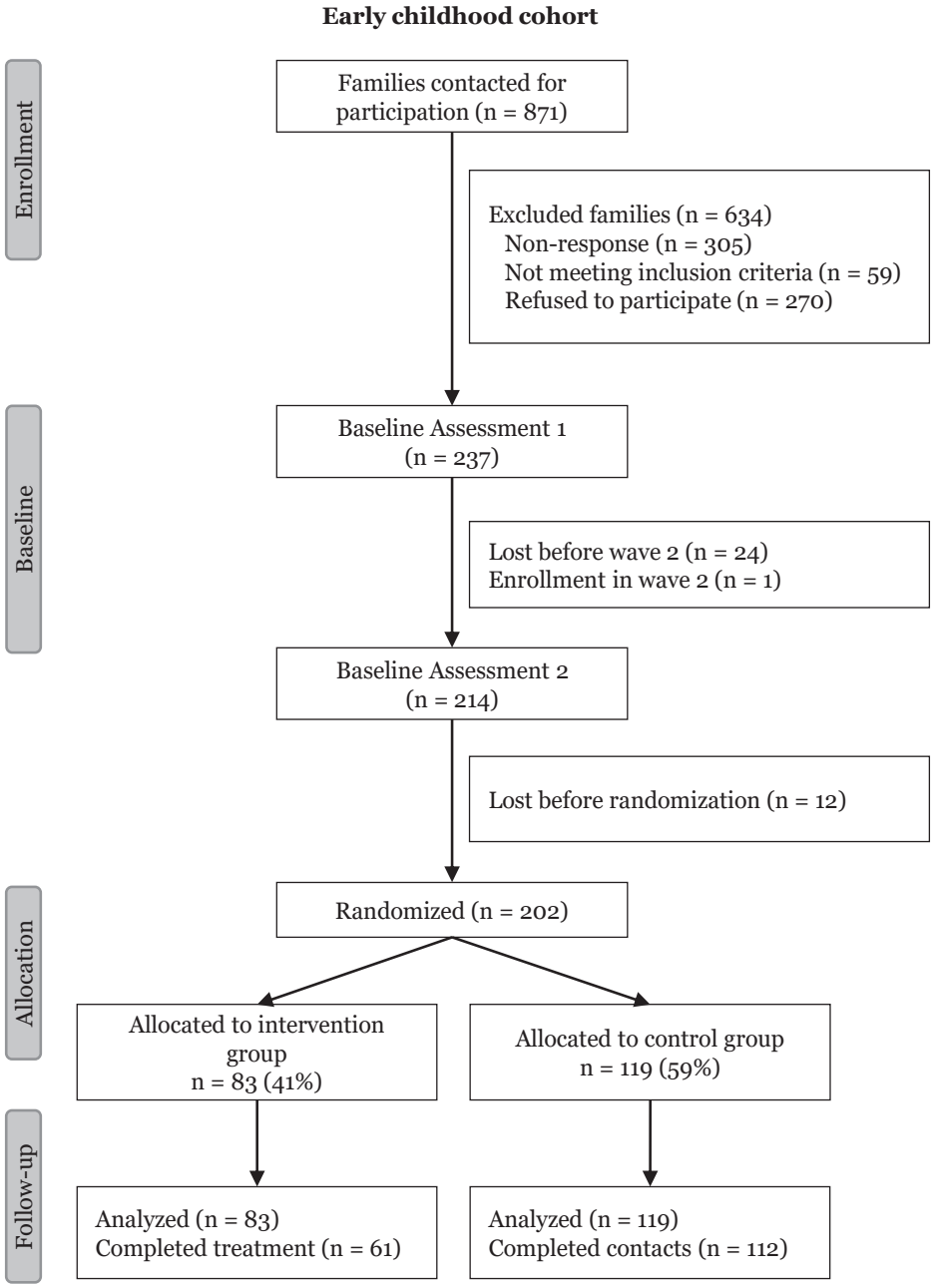


Figure 6.1 continued

Flow chart of the randomized controlled trial with VIPP-SD in both twin cohorts. Number of participants reflects the number of families (including two twin children).

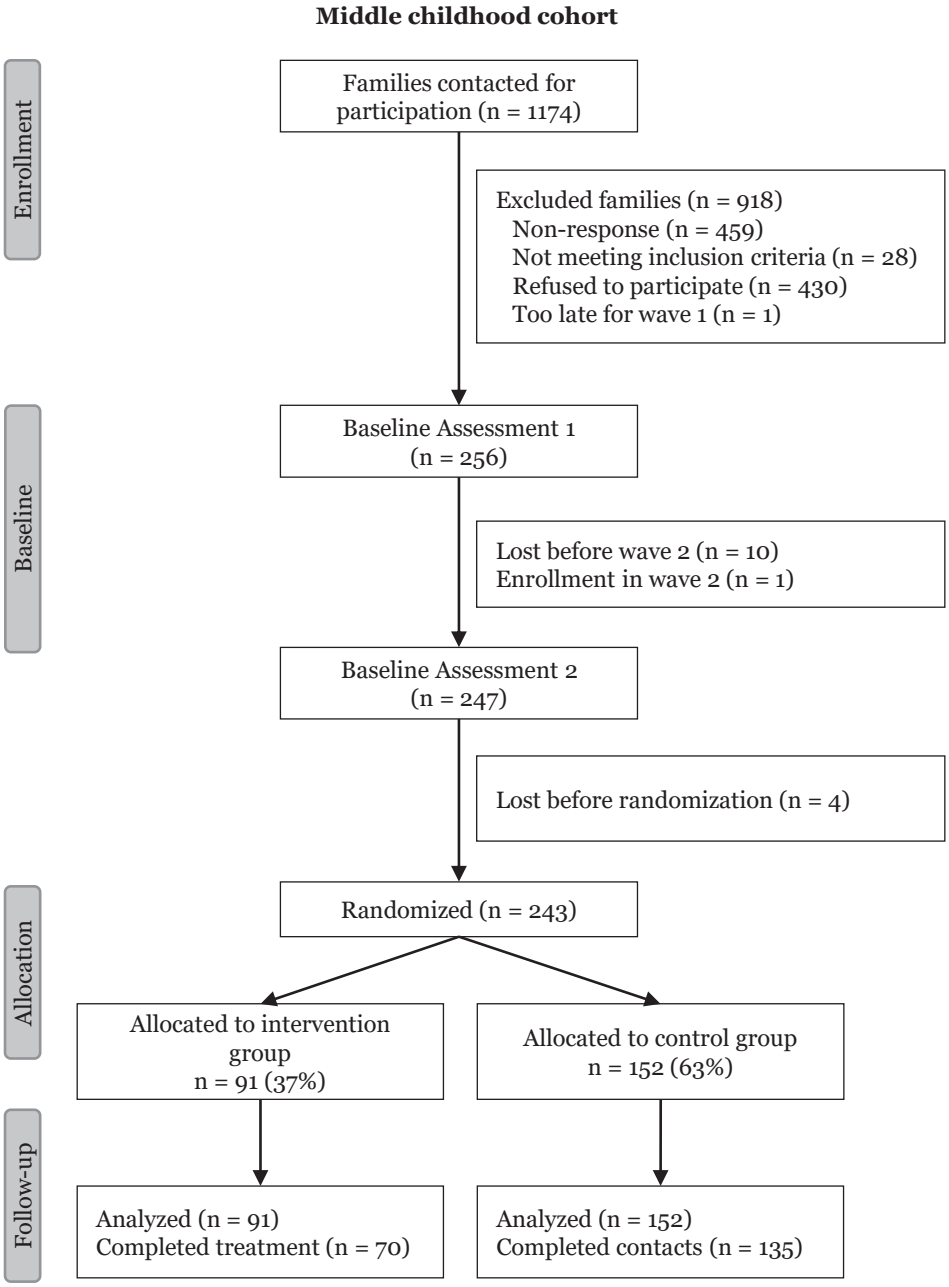


Table 6.1

Characteristics of the complete sample, and separately for the intervention and control groups.

	Early Childhood Cohort			Middle Childhood Cohort		
	Total (n = 202)	Inter-vention Group (n = 83)	Control Group (n = 119)	Total (n = 243)	Inter-vention Group (n = 91)	Control Group (n = 152)
Twin characteristics						
Age M (SD)	3.76 (0.57)	3.74 (0.62)	3.77 (0.53)	7.92 (0.66)	7.94 (0.66)	7.92 (0.67)
Sex (% boys)	45	45.8	44.5	48.6	49.5	48.0
Country of birth (% the Netherlands)	99.5	100	99.2	99.2	100	98.7
Zygosity (% MZ)	60.4	66.3	56.3	55.1	50.5	57.9
Family characteristics						
Primary parent (%)						
Biological mother	91.6	88.0	94.1	90.5	87.9	92.1
Adoptive mother	0	0	0	0.8	1.1	0.7
Biological father	8.4	12.0	5.9	8.6	11.0	7.2
Age primary parent M (SD)	36.87 (4.69)	36.89 (4.82)	36.82 (4.62)	40.48 (4.66)	40.77 (4.78)	40.32 (4.60)
Country of birth (% the Netherlands)	96.0	98.8	94.1	97.5	96.7	98.0
Educational level primary parent						
Lower and Intermediate vocational	30.2	37.3	25.2	34.3	35.2	33.8
Higher vocational, university bachelor	42.1	36.1	46.2	41.7	39.6	43.0
Post-higher vocational, university master	27.7	26.5	28.6	24.0	25.3	23.2
Primary parents' marital status (%)						
Two-parent household	96.5	96.4	96.6	93.8	93.4	94.1
Single parent household	3.5	3.6	3.4	6.2	6.6	5.9

Note. No differences between intervention and control group, if not otherwise reported; measures were taken at T1.

Procedure

Through municipality records, twin families from the western region of the Netherlands were contacted. Families were selected if the twins had the same sex, their parents were Dutch speaking, and parents as well as grandparents were born in Europe. Families with same sex twins were selected to prevent within-twin pair differences due to different sex, which would also lower the power of statistical analyses. The occurrence of congenital disability, psychological disorder, chronic illness, hereditary disease, visual/hearing impairment, or an IQ of <70 led to exclusion. Families received an invitation letter and an information brochure. Parents who indicated a willingness to participate were called to check the inclusion criteria and to provide more information about the study. For a detailed description of the recruitment, see Euser et al. (2016) for the early childhood cohort and Van der Meulen et al. (2018) for the middle childhood cohort. Six yearly visits were scheduled (alternating home and laboratory visits) with additional ambulatory assessments. The current study used data from the first, second, third, and fourth time of measurement (T1, T2, T3, T4) of the early childhood cohort (collected in 2014–2018) and middle childhood cohort (collected in 2015–2019). The intervention took place between T2 and T3. One month after the final intervention or control sessions had taken place, data for T3 were collected.

Ethical approval for the study was provided by the central committee on research involving human subjects (CCMO; Early childhood cohort NL49069.000.14, Middle childhood cohort NL50277.058.14). The study adheres to the CONSORT guidelines (see Appendix). The trial, the study design, and the analysis were pre-registered.

Intervention

Randomization. Using a computer-generated blocked randomization sequence, we randomized the sample at a ratio of 2:3 at the family level stratified by timing of the intervention and twin sex. A ratio of 2:3 was chosen as we had to restrict the number of families receiving the intervention due to limited resources. We randomized the sample after T2 to minimize selective attrition. The researcher who assigned the families to one of both conditions was not involved in data

collection, coding or analysis. Interveners and families were blind to the condition before randomization but not afterwards***du***e to the open label design. Coders and researchers involved in data coding or analysis were blind to the allocation. In total, 174 (39%) families were allocated to the intervention group and 271 (61%) families to the control group.

VIPP-SD for Twins. We implemented the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD, adapted for twin families (Euser et al., 2021) between T2 and T3 in 39% of the families randomly assigned to the intervention group. The VIPP-SD is comprised of five biweekly sessions at the family's home conducted by an intervener. In each session, the intervener first videotaped 15 minutes of parent–child interaction and then gave feedback on the video-recorded interaction of the previous session. During feedback, positive and successful interaction moments were highlighted, and alternatives for insensitive interactions were discussed with the parents. Each session had its own theme (see Juffer et al. (2017) for an overview of the themes). We invited the partner of the primary parent to the final session in order to support the primary parent's implementation of positive parenting behaviors, according to protocol. As we conducted the intervention in twin families, minor adaptations to the original VIPP-SD program were made to address the challenges faced by the parents of twins. These adaptations included issues regarding dividing attention to both children at the same time, and dealing with jealousy and competition among the twin children. Also, toys and games were adapted to the situation of interacting with twins. For a detailed description of the intervention, see Euser et al. (2021) and Runze et al. (2022).

Control Condition. Families in the control condition received five phone calls following a standard protocol parallel to the intervention sessions, to ensure the same number of contact occasions. Using a semi-structured interview, families were asked general and specific questions about their children's development. When parents asked for advice on parenting issues, they were referred to online information or, in case of twin-specific questions, to the Dutch organization for parents with multiples (NVOM).

Measures

Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). Both parents completed the SDQ for both children, a 24-item questionnaire which consists of five subscales, one of which measures conduct problems (five items). Items are scored on a 3-point scale ("not true", "somewhat true" or "certainly true"). An example item of the conduct problems scale is "often fights with other children or bullies them". The reports of both parents correlated significantly (between $r = 0.42$ and $r = 0.55$); therefore, mean scores were computed. Cronbach's alpha ranged between 0.65 and 0.82 across all time points and cohorts (see Table S6.1).

Hair Cortisol Concentrations (HCC). Hair samples were collected by a trained research assistant during the lab or home visit. Collecting hair is a non-invasive method, because only small amounts of hair are needed. Several months of cortisol secretion can be assessed (Rippe et al., 2016). As hair grows approximately 1 cm per month, every 1 cm segment of hair represents the past month. To collect hair samples, a strain of hair at the base of the vertex posterior of the scalp was selected and cut right at the scalp. Hair samples were put into foil and stored at a dark location at room-temperature until sent to the Dresden Lab Service GmbH in Germany for analysis. The most proximal 2 cm of hair were sectioned, representing cortisol production in the past two months. Liquid chromatography–mass spectrometry (LC–MS/MS) was used for cortisol quantification (Gao et al., 2013). The lower limits of quantification were below 0.1 pg/mg. The inter- and intra-assay coefficients of variance were below 10%. For more details regarding the analysis, see Gao et al. (2013).

Genotyping and Imputation. Saliva samples were collected from the children at T2. The DNA was genotyped by the Genetic Laboratory of the Department of Internal Medicine (Population Genomics) at Erasmus MC using the GSA-MD array (version 3). The DNA QC was performed in PLINK (Purcell et al., 2007). After genotyping, the 1000 Genomes Project (phase III release version 5) was used to apply a two-step genotype imputation. The genetic ancestry of the children participating in this study was characterized using the genomic components equivalent to the principal components (PCs) of Europeans (CEU population). The

first 5 PCs of the European-only sample were subsequently used as covariates in our analyses to adjust for spurious population stratification.

Polygenic Score (PGS) of Differential Susceptibility. We computed two different polygenic differential susceptibility scores. We used the GWAS summary statistics of Keers et al. (2016) to obtain a PGS of environmental susceptibility (PGS-ES). For secondary analyses, we used a PGS for the 'susceptibility to environmental stress and adversity' (SESA) cluster, as a potentially relevant construct of differential susceptibility (subsequently called PGS-SESA) (Nagel et al., 2020). For both PGSs we used the PRSice software. The GWAS summary statistics served as the base sample, and L-CID was the target sample. Only autosomal SNPs were used, since there is no consensus for the sex chromosomes (Choi & O'Reilly, 2019). The PGSs were calculated using clump r^2 = 0.1, 250 kb at different p-value thresholds (i.e., 0.20, 0.10, 0.05, 0.01, and 0.001). We tested the PGSs using linear regression models and selected the p-value threshold of the PGSs that explained most variance based on the largest R^2. The PGSs under the best p-value threshold were subsequently used for further analyses (see Table S6.2a,b) (Choi & O'Reilly, 2019). For the PGS-ES, the best p-value threshold for HCC was 0.05 and for conduct problems 0.2. Regarding the PGS-SESA, the best p-value threshold for HCC was 0.05 and for conduct problems was 0.001.

Control Variables. Cortisol levels in hair can be influenced by several environmental factors, which we controlled for. These were sex, BMI, socio-economic status, the number of persons in a household, ethnicity, hair color, last hair wash, and frequency of hair washing (Rippe et al., 2016; Stalder & Kirschbaum, 2012). Although covariate adjustment is not necessarily needed in randomized trials, we controlled for them as this can increase power (Kahan et al., 2014). In the early childhood cohort, children were approximately 5 years old at the time of the intervention, whereas in the middle childhood cohort they were approximately 10 years old at the time of the intervention, so we investigated whether age was a moderator of the intervention effects.

Data Analysis

We conducted intent-to-treat analyses and applied multilevel models with robust full maximum likelihood estimation using MPlus (Muthén & Muthén, 2017). We had three levels (time of assessment, child, family). Each child had four data points for conduct problems and two data points for HCC (repeated measures), and the twin children were nested within families. We computed intraclass correlation coefficients (ICCs) to estimate the proportion of variance explained by the three levels. First, we computed the ICCs by fitting an intercept-only model with three levels and an unrestricted within-subject (co)variance structure separately for conduct problems and HCC. In the next step, we fitted growth models. For conduct problems, time and time^2 were added as fixed and random effects. Condition, Time^2×Condition, Time^2×Condition×PGS-ES and Time^2×Condition×PGS-SESA were included as fixed effects. For HCC, time was added as fixed and random effect. Condition, Time×Condition, Time×Condition×PGS-ES and Time×Condition×PGS-SESA were included as fixed effects. We standardized all predictor variables before inclusion. We conducted the following sensitivity analyses: (1) including only those participants who received all five intervention sessions (81%), (2) including only families in which mothers are the primary caregivers (92%), and (3) for both cohorts separately.

We computed the a priori power using G*Power 3.1 (Faul et al., 2007) based on a repeated measures MANOVA with two outcome variables, an $\alpha = 0.05$ and an effect size of $f = 0.37$, based on the effect size of Poehlmann-Tynan et al. (2020). With a sample size of 798 (both cohorts, sample size at T4), the a priori power was excellent (0.99) for detecting main effects. Even with a small effect size of $f = 0.10$, the power was good (0.81) for detecting main effects. Due to non-convergence issues, we adapted our original pre-registered analysis plans. Please see the Supplementary Materials for a detailed description of deviations from the preregistration.

Results

Preliminary Analyses

Means and standard deviation of the outcome variables can be found in Table 6.2. For hair cortisol, the distributions showed 19 outliers and for conduct problems 20 outliers (*M* +/− >3.29 *SD*), which were winsorized, receiving a value between the highest/lowest non outlying value and the mean +/− 3.29 *SD* (Table S6.3). Of all the possible covariates (see Table S6.4a for descriptive statistics), BMI, the age of the child, the sex of the child, the frequency of washing the hair, and the time of the last hair wash emerged as significant covariates for cortisol (Table S6.4b,c). The inclusion of these covariates, and the five PCs, led to the non-convergence of the model. Therefore, we conducted a regression analysis with the covariates as predictors and cortisol as outcome (R^2 = 14.7%) and used the residuals as the outcome variable in subsequent analyses. There were missing data in both outcome variables (ranging from 2.2% to 17.6%, see Table S6.3). Little's MCAR test was significant (X^2 = 64.91, $p < 0.001$), indicating that data were not missing completely at random. The missing data were subsequently imputed using Expectation-Maximization single imputation. Data were imputed using the EM option in SPSS with a maximum of 25 iterations and intervention group, age, zygosity, and sex as predictors.

Main Results

For conduct problems, the ICC in the intercept-only model with three levels revealed that 11% of the variance in conduct problems could be attributed to the family level, meaning that conduct problems within a family were more similar to each other than conduct problems between families. Less than 1% percent could be attributed to the child level, meaning that conduct problems across time points within a child were similar compared to scores from the twin sibling.

The same was true for HCC problems: the ICC in the intercept-only model with three levels revealed that 23% of the variance in HCC could be attributed to the family level and less than 1% could be attributed to the child level. In subsequent analyses, we therefore included only family and time as levels.

The first hypothesis was that children in the intervention group would exhibit a (stronger) decrease in conduct problems compared to children in the

control group. The interaction between condition and squared time was significant, suggesting that the intervention reduced child conduct problems (b = −0.07, se = 0.03, p = 0.03, Figure 6.2). Our second hypothesis was that the HCC of the children would be significantly affected by the intervention. The interaction between time and condition was significant, implying that in the intervention group cortisol levels decreased more compared to the control group (b = −0.44, se = 0.19, p = 0.02, Figure 6.3). In the third hypothesis, we suggested that the intervention effects on conduct problems and HCC would be moderated by the genetic differential susceptibility of the child as marked by the PGS-ES and the PGS-SESA. For neither PGS did we find a significant three-way interaction (Time^2×Condition×PGS or Time×Condition×PGS, for conduct problems and HCC, respectively), suggesting that these PGSs did not moderate intervention effects on conduct problems (see Table 6.3) or HCC (see Table 6.4).

Sensitivity Analyses

Our sensitivity analyses revealed the following: (1) When repeating the analyses including only those participants who received all five intervention sessions (82% of the families), we found the same results of the VIPP–SD on conduct problems (b = −0.10, se = 0.04, p < 0.01) and on HCC (b = −0.55, se = 0.16, p < 0.01) as in the main analysis (see Table S6.5a,b). (2) In the analysis including only participants where the mother was the primary caregiver (91% of the families), results were comparable to the main analyses, although the interaction between time and condition was statistically significant only for conduct problems (b = −0.08, se = 0.04, p = 0.03) and fell just short of significance for HCC (b = −0.39, se = 0.22, p = 0.07) when using p < .05 as a cut-off point (see Table S6.a,b). (3) The sensitivity analysis for both cohorts separately revealed that, in the early childhood cohort, the Condition×Time^2 effect on conduct problems (b = −0.06, se = 0.05, p = 0.26), and HCC (b = −0.34, se = 0.25, p = 0.16) was not significant anymore. In middle childhood, we found the same results of the VIPP–SD on conduct problems (b = −0.09, se = 0.04, p = 0.03) and HCC (b = −0.71, se = 0.22, p < 0.01) as in the main analysis (see Table S6.6a,b for the estimates).

Table 6.2

Means And Standard deviation of the outcome variables.

	T1	T2	T3	T4
	M (SD)	**M (SD)**	**M (SD)**	**M (SD)**
Hair cortisol (HCC) in pg/mg				
Intervention group	na	3.25 (4.09)	3.04 (4.60)	na
Control group	na	3.08 (3.51)	3.04 (4.47)	na
Conduct problems				
Intervention group	1.30 (0.27)	1.29 (0.27)	1.26 (0.26)	1.21 (0.23)
Control group	1.30 (0.28)	1.29 (0.29)	1.26 (0.26)	1.23 (0.26)

Note. HCC was only collected at T2 and T3.

Figure 6.2

Effect of the VIPP-Intervention on conduct problems.

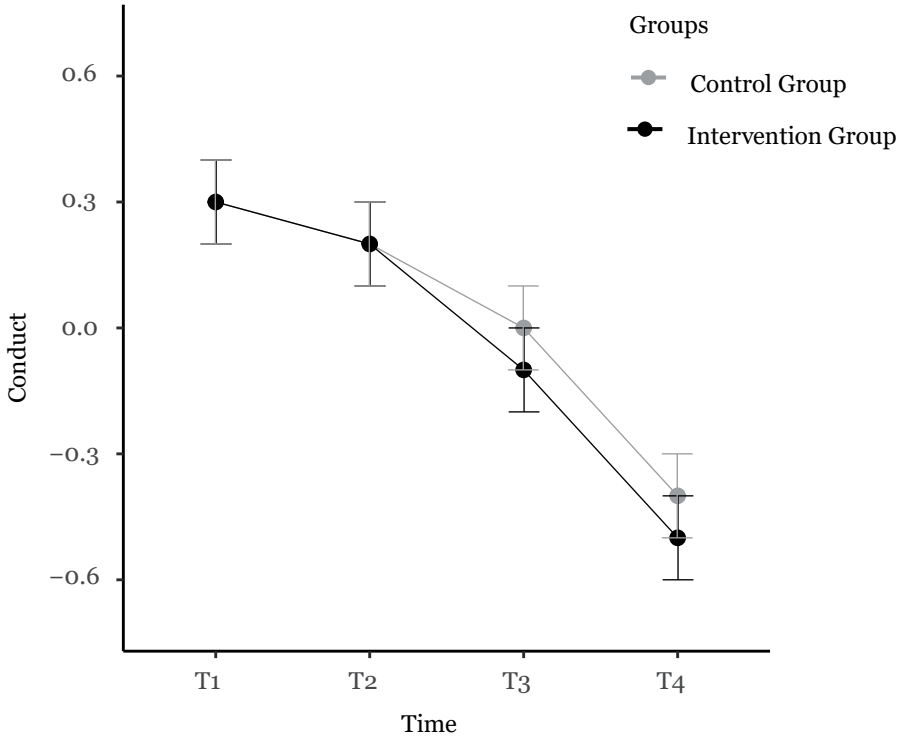


Figure 6.3

Effect of the VIPP-Intervention on hair cortisol levels

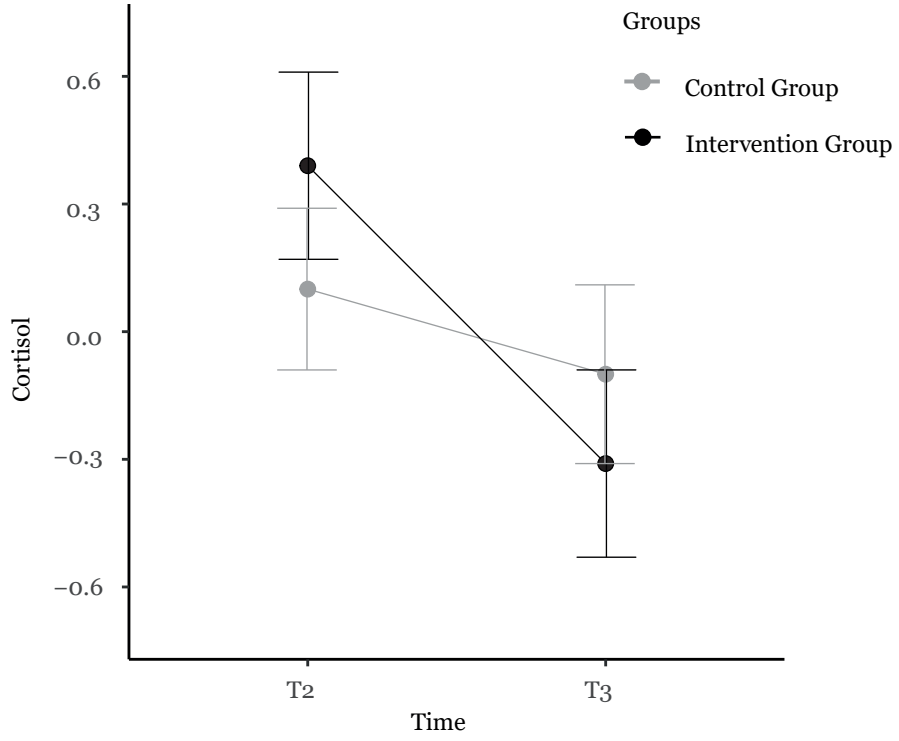


Table 6.3

Multilevel model statistics testing the intervention effect and moderator effect on conduct problems.

Predictor	Est	SE	p	95% CIs
Intercept	0.05	0.07	0.74	-0.06 – 0.16
Time	0.09	0.10	0.33	-0.06 – 0.25
Time2	-0.24	0.10	0.01	-0.38 – -0.08
Condition	0.02	0.06	0.27	-0.08 – 0.11
Condition×Time²	-0.08	0.04	0.03	-0.13 – -0.02
Condition×PGS-ES×Time²	-0.00	0.00	0.41	-0.11 – 0.04
Condition×PGS-SESA×Time²	-0.05	0.04	0.20	-0.09 – 0.01

Table 6.4

Multilevel model statistics testing the intervention effect and moderator effect on hair cortisol

Predictor	Est	SE	p	95% CIs
Intercept	0.09	0.09	0.32	-0.06 – 0.23
Time	-0.12	0.05	0.02	-0.20 – -0.04
Time2	0.50	0.19	0.01	0.20 – 0.81
Condition	-0.44	0.19	0.02	-0.75 – -0.13
Condition×Time^2	0.06	0.06	0.34	-0.04 – 0.15
Condition×PGS-ES×Time^2	0.00	0.06	0.95	-0.09 – 0.10
Condition×PGS-SESA×Time^2	-0.05	0.04	0.20	-0.09 – 0.01

Discussion

In the current study, we investigated whether a brief interaction-focused parenting intervention, the VIPP-SD, had an effect on children's conduct problems and hair cortisol concentrations. In addition, we examined whether the intervention effect was moderated by children's differential susceptibility captured by two novel polygenic scores. The VIPP-SD was successful in decreasing children's conduct problems in our population-based sample. This adds new evidence to the meta-analysis of VIPP-SD studies that reported an overall non-significant effect on child externalizing behaviors (van IJzendoorn et al., 2023). In that meta-analysis, only nine studies reported on externalizing behaviors, and the mean sample size of the nine studies was 116, with only three studies including more than 100 participants. It may thus have been underpowered for detecting the effects of the strength that we found in this study. Indeed, the most recent pre-registered and best-evidence study with 300 participants did find a significant effect of VIPP-SD on externalizing behaviors (O'Farrelly et al., 2021) with an effect size that was similar to the effect found in our study.

In the intervention group, children's hair cortisol concentrations decreased

more after the intervention compared to the control group, indicating that the VIPP-SD was successful in decreasing hair cortisol levels. Our findings are in line with the only other published intervention study that investigated child hair cortisol concentrations as an outcome of a randomized trial. Poehlmann-Tynan et al. (2020) found that after cognition-based compassion training (CBCT) with 25 parents and their children between 4 months and 5 years, child hair cortisol levels were significantly lower than those of children in the waitlist control group. We replicated these results in a well-powered twin study, implementing a parenting intervention based on attachment theory and the social learning model of coercive cycles (Bosmans et al., 2022; Juffer et al., 2017). A recent meta-analysis of 19 parenting intervention studies found no significant effect on salivary cortisol (Martins et al., 2020). The difference may be explained by the fact that salivary cortisol provides a snapshot of cortisol levels at a specific time or day, and cortisol levels are known to fluctuate substantially during and across days (Stalder & Kirschbaum, 2012). Hair cortisol is thought to provide a more stable picture of cortisol secretion across time, and may thus indicate chronic stress levels instead of more volatile, temporary states of stress exposure or, e.g., physical exercise. However, hair cortisol is subject to confounding factors such as hair washing frequency, which can introduce a large amount of measurement error (Stalder et al., 2017). In the case of uncontrolled confounding, the verdict is still open on which cortisol measurement method is the better choice.

We tested differential susceptibility to the positive effects of the intervention using two polygenic scores. We did not find support for differential susceptibility to parenting concerning conduct problems or hair cortisol concentrations. One explanation might be that susceptibility might not be domain-general but domain-specific (Boyce & Ellis, 2005; Ellis et al., 2011). For example, Zhang et al. (2023) found that individual differences in susceptibility to family–social effects were not related to individual differences in susceptibility to quality-of-care cognition-related effects, indicating that individuals who were susceptible to family–social effects were not susceptible to quality-of-care effects and vice versa. However, they did find that individual differences in susceptibility to family and child-care effects were positively correlated. Considering the similarity of family effects (care environment at home) and child care effects (care environment at child care), the significant correlation

between susceptibilities does not inevitably point to a domain generality as the authors suggested but rather to a broader but specific domain (caregivers). Indeed, a recent paper by Belsky et al. (2022) found that approximately 50% of the children who were highly susceptible to the effect of childcare quality on pre-academic skills were not highly susceptible to the effect of child care quantity on behavior problems and vice versa.

We speculate that the two PGSs in our study might be indicators of differential susceptibility, but specific to the outcome domains for which they were developed, and not for the specific exposure or outcome domain we focused on in this parenting intervention study. The PGS-ES is a polygenic score for differential susceptibility based on GWAS data for anxiogenic environments (Keers et al., 2016). Keers et al. (2016) found that the PGS-ES moderated the association between child-reported parenting at age 12 and children's internalizing emotional problems. Lemery-Chalfant et al. (2018) found that the PGS-ES moderated the association between the effects of the Family Check-Up intervention and internalizing psychopathology. In the current study, we considered the possibility that susceptibility towards the effects of parenting on internalizing and externalizing behaviors might be indicated by one underlying PGS, but we were not able to corroborate this assumption. Further research is needed to investigate whether this PGS-ES is a susceptibility marker for any outcome outside the internalizing domain.

We were the first to explore the PGS-SESA, a PGS based on the GWAS of susceptibility to the environmental stress and adversity (SESA) cluster of neuroticism (Nagel et al., 2020), as a moderator of the effects of a parenting intervention on children's conduct problems and hair cortisol concentrations. Considering the possibility of domain-specificity of differential susceptibility, we can only cautiously conclude that in our study the PGS-SESA did not emerge as a differential susceptibility marker. Susceptibility to environmental stress and adversity in the context of neuroticism might be more closely related to the perception of stress and adversity than to the neurobiological experience of stress that is central to the Boyce and Ellis model of sensitivity to stress (Boyce & Ellis, 2005; Ellis et al., 2011). Another issue might be the age difference between the sample providing the PGS-SESA and the children in our study. Although DNA does not change across age, its

interplay with environmental influences might fluctuate across time (Slagt et al., 2016; Windhorst et al., 2015). More research on the PGS-SESA is clearly required to draw firm conclusions about its role as a marker of differential susceptibility to various environments, in various ages, and for different outcomes.

Some limitations should be noted. We measured child conduct problems using parent-report questionnaires. We cannot exclude the possibility that the intervention increased parents' appreciation or tolerance of child behaviors. Parents might have reported fewer conduct problems after the intervention because of a change in the way they interpreted their child's behavior. However, we included the reports of both the primary parent who participated in the intervention and the other parent who was only invited to one booster session. It should also be noted that the intervention was successful in improving parental sensitive discipline in the early childhood cohort (Euser et al., 2021), whereas in the middle childhood cohort parents' attitudes on sensitive parenting improved, but not their behavior one month after the intervention (Runze et al., 2022). As reported in Runze et al. (2022), the task within which sensitive discipline was measured in the middle childhood cohort might have been too complex. The intervention may have had an effect on a dimension of parenting behavior that was not measured at posttest but was relevant for its influence on child behavior problems. As imputation strategy, we used single imputation with the Expectation Maximization, where it is not possible to include interactions as predictors. Therefore, there may be bias in the standard errors and the estimation of the interactions. Another limitation concerns the skewed distribution of conduct problem scores in our low-risk sample, decreasing the power to detect (moderated) differences between the intervention and control groups. A better distribution of the outcome measure or a much larger sample size would increase power and would also help to facilitate model convergence. Gene × Environment (G × E) interactions have been criticized because of a basic lack of statistical power similar to the test of any interaction term outside the realm of genetics but this critique focused on correlational designs (Duncan & Keller, 2011). In experimental designs, the statistical power of tests for (G × E) interactions is much larger (Bakermans-Kranenburg & van IJzendoorn, 2015).

Having said that, we believe that our study provides a sufficiently firm basis

for the conclusions drawn. We collected data in a relatively large sample and included four waves of data collection for conduct problems. Together with the preregistration and the RCT design, this gives us confidence in the replicability of our study results. Bearing in mind that most phenotypes are influenced by many genetic variants (Moffitt et al., 2006), we used two novel polygenic scores to test for differential susceptibility as opposed to single candidate genes as has been done in the past. The current study is among the first set of randomized trials to use some GWAS-based polygenic scores as markers for differential susceptibility. Although we could not confirm their role as moderators of intervention effects, this approach is an important first step to the identification of the possible genetic origins of differential susceptibility in a domain with rather complex exposures as well as outcomes.

Conclusion

In this randomized controlled trial with 445 families from two twin cohorts including 890 children, we found that the Video Feedback Intervention to Promote Positive Parenting and Sensitive Discipline (VIPP-SD) was successful in decreasing conduct problems and hair cortisol concentrations in school-aged children. Intervention effects were not moderated by the two polygenic scores specifically developed for other domains and other ages, suggesting the potential age- and domain-specificity of differential susceptibility.

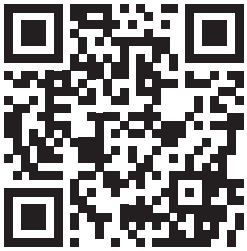
Acknowledgments

We thank the participating families for their enthusiastic involvement in the Leiden Consortium on Individual Development (L-CID). We are also grateful to the data-collection and data-processing team, including all current and former students, research assistants, students, and postdoctoral researchers for their dedicated and invaluable contributions. Marinus van IJzendoorn, Eveline Crone and Marian Bakermans-Kranenburg designed the L-CID experimental cohort-sequential twin study “Samen Uniek” as part of the Consortium on Individual Development (CID; Gravitation grant 2013–2023 awarded by the Dutch Ministry of Education, Culture, & Science, and the Dutch Research Council, NWO grant number 024.001.003).

Supplementary Materials

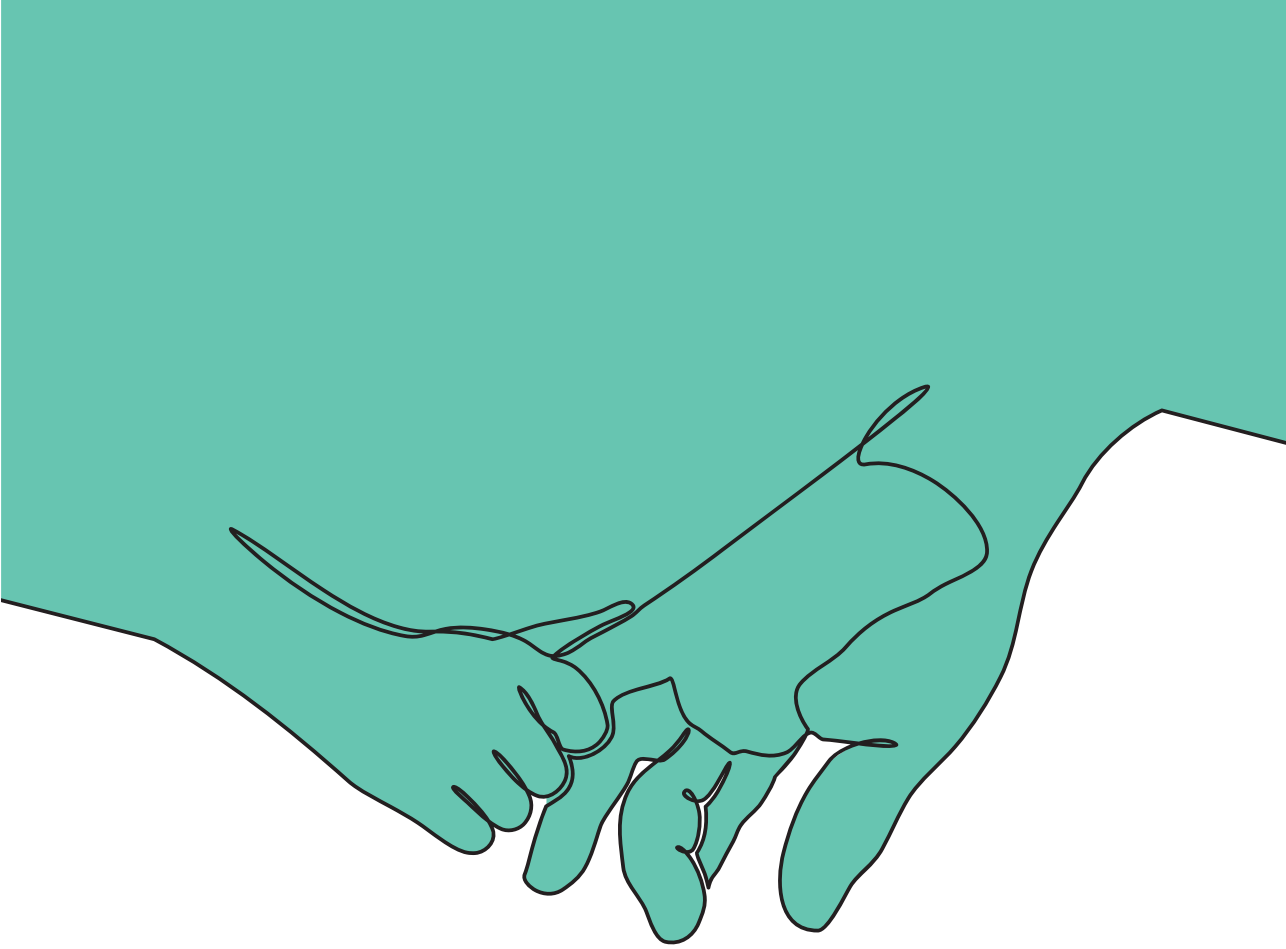
All supplementary materials are published and can be retrieved from:

http://tinyurl.com/Chapter6Supplement



This chapter is under review as:

Runze, J., Witte, A.M., Van IJzendoorn, M.H., Oosterman, M., & Bakermans-Kranenburg, M.J. (2023). Differential Susceptibility in the Intergenerational Transmission of Secure Base Script Knowledge? *Under review*



DIFFERENTIAL SUSCEPTIBILITY IN THE INTERGENERATIONAL TRANSMISSIONT OF SECURE BASE SCRIPT KNOWLEDGE?

CHAPTER SEVEN

Abstract

Verhage et al. (2016) documented evidence supporting the transmission of attachment from parents to their children, with parental sensitivity serving a mediating role. Nevertheless, a "transmission gap" exists. In the current pre-registered study, we investigated whether parents' secure base script knowledge predicted their children's secure base script knowledge and whether this association was mediated by parental sensitivity and sensitive discipline. Furthermore, we examined whether these associations were moderated by the child's differential susceptibility. We used the Attachment Script Assessment (ASA, Waters & Waters, 2021) with parents and their 9-year-old children in one cohort (N = 239 families) of the L-CID longitudinal twin study. We used a polygenic score and child discomfort as differential susceptibility factors in the intergenerational transmission of attachment. Multilevel structural equation models show that parental sensitivity, but not parental secure base script knowledge or parental sensitive discipline predict children's attachment. This association was moderated by child temperamental discomfort: lower levels of parental sensitivity predicted less secure child attachment in children with higher discomfort, but not in children with lower discomfort. If replicated, we may conclude that the intergenerational transmission of secure base script knowledge is moderated by temperament in a way consistent with the diathesis-stress model.

In the long-standing research tradition of attachment research, the transmission of attachment from parent to child and the lack of transmission of attachment in some cases is a well-known phenomenon. In a comprehensive meta-analysis comprising 95 studies, Verhage et al. (2016), like their predecessor Van IJzendoorn (1995) who analyzed 19 studies two decades earlier, documented evidence supporting the transmission of parental attachment representations (as measured with the Adult Attachment Interview, Main et al., 1985) to their children's attachment behavior (as measured with the Strange Situation Procedure, Ainsworth, 1978), with parental sensitivity serving as a partial explanatory factor. Nevertheless, despite the extensive examination of these studies, the authors concluded that a "transmission gap" persisted of the correspondence between parental attachment representations and child attachment that is not explained by parental sensitivity. Other parental constructs might explain part of this gap, such as parental sensitive discipline, which has been found to be predicted by parental secure base script knowledge (Witte et. al., 2023). In the current study we aimed to test observed parental sensitivity as well as parental sensitive discipline as mediators in the intergenerational transmission of attachment as indicated by parents' and their children's secure base script knowledge. As a further aim, we drew inspiration from Verhage et al.'s (2016, p. 31) speculation that "some children might be more resilient against negative influences", and investigated whether children's differential susceptibility to parenting may serve as a moderator explaining part of the transmission gap.

Over time, different measures of attachment in adults and children have been developed and used in attachment research, such as the Adult Attachment Interview (Hesse, 2016; Main et al., 1985) and, in children, the Attachment Q-sort (AQS; Waters & Deane, 1985) and the Strange Situation Procedure (Ainsworth, 1978; Ainsworth et al., 2015). Another measure is the Attachment Script Assessment (ASA), developed by Waters and Waters (2006). This measure is based on the notion that experiences with attachment figures are organized into a cognitive script: the secure base script (Waters & Waters, 2006; Waters et al., 2021). Waters and Waters (2006) note that individuals who received consistent and sensitive caregiving are more likely to develop comprehensive and readily accessible secure base scripts. These scripts enable individuals to rely on their attachment figures for effective

support during moments of distress. In contrast, individuals who have experienced inconsistent, insensitive, or atypical caregiving are expected to have fragmented, inconsistent or ineffective secure base script knowledge (SBSK, Nivison et al., 2021; Waters & Waters, 2006; Waters et al., 2021). The measure is particularly fit for adults and children over age six, has been found to be reliable and valid, and is relatively easy to administer (Waters et al., 2019; Waters et al., 2015).

The secure base script is based on eight central elements which represent a temporal-causal sequence to elicit effective secure base support: (1) The individual engages constructively with the environment. (2) The individual encounters a threat that prevents constructive engagement and/or results in distress. (3) The individual signals the need for support. (4) The partner individual (e.g., the parent) adequately picks up this signal and provides support, which is (5) accepted and (6) is effective in resolving the threat. (7) The support provided effectively relieves the distress and (8) the individual can return to engage constructively with the environment (Waters & Waters, 2006). Previous research indicated that SBSK is moderately to strongly associated with coherence scores and security scores on the Adult Attachment interview (Coppola et al., 2006; Dykas et al., 2006; Hawkins et al., 2015; Steele et al., 2014), is relatively stable over time (Vaughn et al., 2006; Waters et al., 2017, 2021, 2022), and can be changed through attachment-related experiences (Bosmans et al., 2020, 2022; Waters et al., 2019). As parental sensitivity has been found to predict SBSK in children (Nivison et al., 2021; Steele et al., 2014; Vaughn et al., 2016; Waters et al., 2017) and the ASA is suited for children from age six onwards, it is an excellent measure for the investigation of the transmission of attachment representations from parents to school-aged children.

The differential susceptibility theory posits that children exhibit variability in how much they are affected by a positive or negative environment, including parental behavior (Belsky et al., 2007; Ellis et al., 2011). Highly susceptible children are expected to display more psychological problems when exposed to adverse environments. However, the same children may also benefit above average from supportive environments. In contrast, less susceptible children are thought to be less affected by both adverse and supportive environments. Previous research has used three groups of markers of differential susceptibility. These include genetic variables (such as

dopamine-related and serotonin-related polymorphisms, Bakermans-Kranenburg & Van IJzendoorn, 2015; Belsky & Van IJzendoorn, 2017), endophenotypic variables (such as cardiovascular stress reactivity or sympathetic nervous system reactivity or HPA-axis reactivity, Boyce & Ellis, 2005), and dimensions of temperament, in particular negative reactivity, impulsivity, anger proneness, withdrawal negativity or sensitivity (Belsky et al., 2007; Slagt et al., 2016). A promising approach to studying genetic differential susceptibility involves the use of polygenic scores (PGS) obtained from genome-wide association studies (GWAS). These scores serve as a composite marker of differential susceptibility. Keers and colleagues (2016) were the first to develop a polygenic score for differential susceptibility based on GWAS data for anxiogenic environments.

Previous research investigating differential susceptibility using different indicators has produced mixed evidence. For example, Cassidy et al. (2011) examined whether the effect of a brief intervention to increase secure infant attachment was moderated by infant irritability and found that highly irritable infants benefitted more from the intervention compared to moderately irritable infants, but highly irritable infants in the control group did not have significantly less secure attachment bonds compared to moderately irritable infants. Using a different intervention, Video-feedback Intervention to promote Positive Parenting (VIPP), Klein Velderman et al. (2006) found that the highly reactive infants were more susceptible to changes in their mother's sensitivity compared to less reactive infants. In one of our own studies, we were unable to detect differential susceptibility to the influence of an attachment-based intervention on children's hair cortisol levels and behavior problems (Runze et al., 2023). One explanation for the mixed findings and a reason to further investigate the differential susceptibility hypothesis is its theorized domain-specificity (Belsky et al., 2022; Zhang et al., 2023). Individuals are thought to be differentially susceptible to certain environmental factors but not all environments in general. Consequently, null findings do not necessarily mean the absence of differential susceptibility, but rather no differential susceptibility to the specific environment with respect to the specific outcome. Considering the potential specificity, more research on differential susceptibility in the context of parenting and child attachment is needed.

In sum, in the current study we examined the intergenerational transmission

of attachment model (Van IJzendoorn, 1995; Verhage et al., 2016) using the ASA with both parent and child (Waters & Waters, 2006; Waters et al., 2019). We investigated as mediators parental sensitivity, which has consistently been found to play a role in the transmission of attachment, and, additionally, parental sensitive discipline, which we have found to be also predicted by parental secure base script knowledge in an earlier study (Witte et al., 2023) and which might serve to further narrow the transmission gap. Moreover, we investigated the role of child genetic and temperamental differential susceptibility in the intergenerational transmission of attachment.

Methods

Design

The L-CID project is an experimental longitudinal twin study with two cohorts: and early childhood cohort and a middle childhood cohort (see website: www.developmentmatters.nl). Through municipality records, twin families from the Netherlands were contacted if the twins had the same gender. They received an invitation letter and an information brochure on the study. Families were eligible if the parents were Dutch speaking and parents as well as both sets of grandparents were of European descent. The occurrence of congenital disability, psychological disorder, chronic illness, hereditary disease, visual/hearing impairment, or an IQ of <70 led to exclusion. For a detailed description of the recruitment, see Euser et al., 2006. Six yearly visits were planned (alternating home and laboratory visits). In the current study, we used data from the early childhood cohort from the second, third and sixth wave (T2, T3, and T6, see Table S7.1).

Participants

Two hundred and thirty-nine families with twins were enrolled in the study (N = 478 children, 58% monozygotic pairs, 51% female). At T2, the children were on average 4.77 years old (SD = 0.58), and at T6 they were on average 9.09 years old (SD = 0.61). The primary caregivers were on average 37.24 years old (SD = 4.70) and in 92% of the families the primary caregivers were female. In 91% of the families, both parents were born in the Netherlands and 93% of the families were from a

middle or high socio-economic background (based on parental education).

Measures

Parental Secure Base Script Knowledge (SBSK-P). We measured parents' SBSK with the Attachment Script Assessment at T2 (ASA; Waters & Waters, 2006). The ASA consists of six stories, two neutral stories and four attachment stories. To minimize participant burden, we used one neutral story and three attachment stories, as three attachment stories have also reliably been used in other studies (Cuyvers et al., 2023; Verhees et al., 2021; Waters et al., 2015; Waters et al., 2019). Moreover, the ASA has been assessed with high reliability coefficients even using only two stories (Waters et al., 2019). Parents were asked to tell four stories using four prompt word sets, each including a title and a list of 12 prompt words. The prompt words implied a beginning, middle, and end of a possible story. The first prompt word set with an implied neutral story was used for practice purposes. Of the three attachment-related word prompts, two word sets concerned hypothetical mother-child relationships (Baby's Morning and Doctor's Office) and one of the word sets concerned a hypothetical adult–adult relationship (The Accident). The stories were scored on a 7-point rating scale, ranging from 1 = "poorly developed secure base script" to 7 = "well-developed secure base script". All stories were independently rated by two coders and for each narrative we averaged the scores assigned by the two coders. As coders were blind to which story belonged to which parent, they were allowed to code multiple stories from the same parent. When scores for the same narrative deviated more than one point, they were discussed to consensus. Mean scores of the three attachment-related narratives were computed as an index of parents' SBSK. The mean ICC with the expert coder and among coders was (n = 4) was between .73 and .75 based on 120 stories in the reliability set, comparable to ICCs in earlier research (Waters et al., 2015; Waters et al., 2019). Cronbach's alpha for the stories was somewhat lower than in previous research (e.g. α = .70 in Waters et al., 2015) but acceptable (α = .65). See Witte et al. (2023) for a more detailed description of the task.

Child Secure base script knowledge (SBSK-C). SBSK was assessed using the Middle Childhood Attachment Script Assessment (MC ASA) at T6 (Waters et

al., 2015; Waters et al., 2019). Like in the parent version, children were asked to tell four stories (one neutral, three attachment eliciting stories, namely "scary dog in the yard", "at the beach" and "hockey game") based on four prompt word sets with a title and 12 prompt-words or statements. In the original version, one story was about a soccer game which we adapted to a field hockey game as field hockey is played by girls and boys in the country of this study, the Netherlands, whereas soccer is predominantly played by boys in the Netherlands. The prompt-words and statements were unchanged, only the title was changed from "Soccer" to "Hockey". The prompt-words imply a beginning, middle and end of a possible story. Narratives were double coded by six trained coders. When scores for a narrative deviated more than one point, a consensus score was assigned. The mean ICC with the expert coder and among coders was between .80 and .81 on average, based on 45 stories in the reliability set, comparable to ICCs in earlier research (Waters et al., 2015). Scores of the three narratives of both coders were averaged into an overall score of SBSK of the children. Reliability was higher than in previous research (α = .81 compared to α = .78 in Waters et al., 2015).

Observed Parental Sensitivity. We observed sensitivity of the primary parent for each twin separately using a computerized version of the Etch-A-Sketch task (Euser et al., 2020). The order of the observations (oldest child first or youngest child first) was randomly chosen. The parent-child couple made three printed example drawings (with increasing difficulty) on a computer screen (Euser et al., 2021). The interaction between parent and child was filmed and the screen with the drawings was recorded. We coded sensitivity using the revised (Egeland, 1990) 7-point rating scales in concordance with Euser et al. (2021). The scales comprise Supportive Presence (1 = parent completely fails to be supportive to the child, 7 = parent skillfully provides support throughout the session) and Intrusiveness (1 = parent allows the child sufficient time to explore and to attempt to solve tools on her/his own, 7 = parent is highly intrusive; her/his agenda clearly has precedence over the child's wishes). The scales correlated substantially (r = .57), therefore we computed a mean score of the supportive presence scale and the reversed coded intrusiveness scale as done in previous research (Kok et al., 2015). Six coders, trained by two expert coders (MBK and SE), coded the videos. Intercoder reliability

(intraclass correlation coefficient; ICC) with the expert coder and among coders was adequate (.71 - .76).

Observed Parental Sensitive Discipline. We observed sensitive discipline of the primary parent using an adapted version of the Do-Don't task (Euser et al., 2021; Van Der Mark et al., 2002). The order of children completing this task was randomly chosen. The parents received written instructions on it and a bag of toys. Parents had to set limits in two different ways: ensuring that the child would engage in the not so attractive "do" task and preventing the child from engaging in the attractive "don't" task. The "do" task consisted of the parent asking the child to put away attractive toys which were in front of them and with which the child had been playing before. The "don't" task consisted of the children refraining from touching any of the attractive toys for two minutes. After two minutes, the child was allowed to play with the least attractive toy only for another two minutes. Both tasks were filmed. We used the positive discipline scale to measure sensitive discipline (Euser et al., 2021; Runze et al., 2022). Positive discipline was rated on an adapted version of the revised Erickson 7-point rating scale (Egeland, 1990) for supportive presence (1 = parent completely fails to provide positive discipline, 7 = parent skillfully provides positive discipline throughout the session). Seven coders, trained by an expert coder, coded the videos. Intercoder reliability (intraclass correlation coefficient; ICC) with the expert coder and among coders was adequate (.79 - .81).

Differential susceptibility factors. Polygenic score of Differential Susceptibility. Based on within-pair variability in emotional problems in monozygotic twins, Keers and colleagues (2016) identified genetic variants of environmental sensitivity and created a PGS, which was also used in a randomized trial by (Shaw et al., 2019). To compute this polygenic score, we collected saliva samples from the children. The Genetic laboratory of the Department of Internal Medicine (Population Genomics) at Erasmus MC genotyped the DNA using the GSA-MD array (version 2). DNA QC was performed in PLINK (Purcell et al., 2007). After genotyping, SNPs were imputed using the 1000 Genomes Project (phase III release version 5). We used genomic components equivalent to the principal components (PCs) of Europeans (CEU population) to characterize the genetic ancestry of the children. We included the first five PCs as a covariate in our analyses, to adjust for spurious

population stratification (Choi et al., 2020). We used PRSice to calculate polygenic scores (Choi et al., 2020). We calculated PGS using a clumping method with a linkage disequilibrium threshold of r2 = 0.1, and a 250kb distance at different p-value thresholds (i.e., 1, .50, .20, .10, .05, .01 and .001). We tested the PGSs using linear regression models and chose the p-value threshold of the PGSs that explained most variance based on the largest R^2. This PGS was subsequently used for further analyses (Choi et al., 2020). As sensitivity analysis, we repeated the analyses using the PGS with a p-value threshold of 1.

Temperamental Differential Susceptibility. The discomfort subscale of the Children's Behavior Questionnaire (CBQ) measures temperamental difficulty and was filled in by the parent at T2. The scale consists of 12 items and is scored on a seven 7-point Likert scale ranging from 1 = extremely untrue to 7 = extremely true. An example item is "My child notices smoothness or roughness". Cronbach's alpha has been satisfactory in a previous study (α = .73 and .67, Rothbart et al., 2001) and was similar in the current study (α = .73).

Covariates. We included group status in the intervention (intervention group vs control group) as a covariate. Furthermore, we tested whether age of the child, sex of the child, verbal IQ score, gender of the parent and socio-economic status were significantly related to our study variables.

Data analysis

The current study was pre-registered. A priori power was computed using the Shiny app MedPower (Kenny, 2017). Assuming effects similar to previous studies (Verhage et al., 2016), a significance level of alpha = .05 and a sample size of 478 children, power for all paths of the mediation model was excellent (> .98). We first computed the ICC to assess whether the data within families was more similar than across families. Then, we employed multilevel structural equation modeling using the package lavaan (Rosseel, 2012) from R (R Core Team, 2017), which makes it possible to leverage data from both twins while accounting for the nested structure of the data. We used full information maximum likelihood (FIML) and the Yuan-Bentler scaled Chi-square estimator with Huber-White covariance adjustment to the standard errors in case of non-normally distributed data. We assessed model fit

using the Comparative Fit Index (CFI; Bentler, 1990), the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973), and Root Mean Square Error of Approximation (RMSEA; Hu & Bentler, 1998). Good model fit is assumed with CFI and TLI values greater than .95 and RMSEA smaller than .08 (Xia & Yang, 2019). Unfortunately, the structural equation models including the PGS and the principal components did not converge. This is not uncommon as cluster size is small (n = 2) and many variables have no variance at the between level (principal components and polygenic scores are the same for monozygotic children). Therefore, we split the sample randomly in twin 1 and twin 2 to overcome the nested structure and conducted regression analyses including interaction terms for the moderation analysis with the polygenic score. In order to test whether a potential moderation reflects differential susceptibility, we used the online application designed by R.C. Fraley as a supplement to Roisman et al. (2012; https://www.yourpersonality.net/interaction/) and randomly chose one child per family to ensure a non-nested dataset. Ethical approval for the study was provided by the central committee on research involving human subjects (CCMO; NL49069.000.14).

Results

Preliminary results

Table 7.1 shows descriptive statistics, outliers, and percentage of missingness. Table 7.2 shows correlations between the study variables. Outliers (data points that deviated 3.29 SD or more from the mean) were winsorized. We examined missingness using the mcar test of the naniar package (Tierney et al., 2022). Missingness was not completely at random as indicated by Little's MCAR test (X^2 (789) = 1503.47, p < .001), but participants who dropped out during the study did not differ from participants who were included in all waves with regard to child age, sex, psychiatric diagnoses or IQ (Dobbelaar et al., 2023). We imputed the data using multiple imputation which provides less biased estimates in cases of non-MCAR data (Madley-Dowd et al., 2019; Newman, 2014). Using the mice package (Van Buuren & Groothuis-Oudshoorn, 2011), we imputed 40 datasets with 20 iterations with predictive mean matching for all study variables. Age and sex of the child as well

as sex of the parent were included in the imputation model.

We assessed whether covariates were significantly associated with our outcome variable using regression analysis. Only sex of the child emerged as significant covariate and was thus, together with intervention group status, included in the analyses (see Table S7.2). We assessed whether the nested structure of the data explained any variance by computing the ICC using the misty package (Yanagida, 2023). The ICC was .30, which is substantial, therefore we employed multilevel structural equation modelling (Musca et al., 2011).

Intergenerational transmission of Secure Base Script Knowledge

There was no significant association between the SBSK of the parent and the SBSK of the child (β = -.03, p = .461) indicating that there was no direct transmission of attachment representation from parent to child. Equivalence testing revealed that this effect was equivalent to zero (p > .001 given equivalence bounds of ± .21 which corresponds to the lower bound of the confidence intervals of r = .24 found in Verhage et al. (2016)). However, we did find a positive association between parental sensitivity and child SBSK (β = .09, p = .027) with higher sensitive parenting being related to higher SBSK scores. Parental sensitive discipline did not predict child SBSK (β = .04, p = .413).

Differential susceptibility in the transmission of Secure Base Script Knowledge

The association between parents' and child's attachment representations was not moderated by the PGS (child 1: β = -.02, p = .573; child 2: β = -.01, p = .738), nor by child discomfort (β = -.06, p = .087). The PGS did also not moderate the association between sensitivity or sensitive discipline and SBSK of the child (child 1: sensitivity: β = -.01, p = .699, sensitive discipline: β = .02, p = .618; child 2: sensitivity: β = .03, p = .249, sensitive discipline: β = -.03, p = .366, see Table S7.3 and S7.4). Child discomfort moderated the association between parental sensitivity and SBSK of the child (β = .08, p = .005, Figure 1). Regions of Significance (RoS) analysis revealed that in children with high discomfort the positive association between parental sensitivity and child attachment representation was significant (β = .08, t(473) = 2.62, p < .009, based on ± 1 SD), but in children with low discomfort the association

Table 7.1

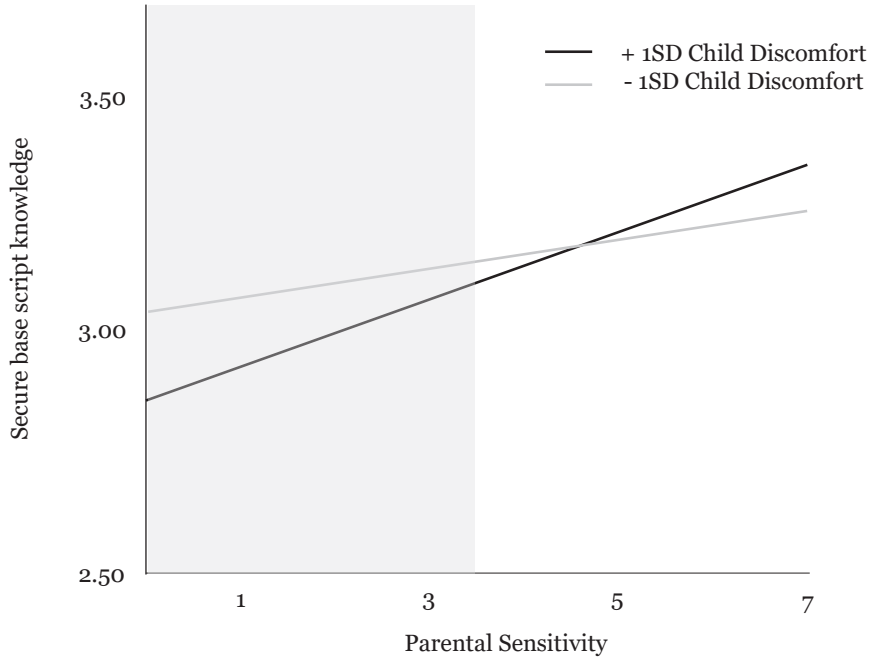
Descriptive statistics of all study variables

Variable	N	% missing	Nr outliers	M	SD	Min	Max	Skew
SBSK-P	400	16%	4	4.02	0.64	1.75	6.00	0.48
Sensitivity	400	16%	0	4.21	1.37	1.00	7.00	-0.11
Sensitive Discipline	400	16%	0	4.95	1.31	1.50	7.00	-0.33
SBSK-C	313	34%	2	3.19	0.47	1.00	4.67	-0.92
Discomfort	396	17%	2	3.64	1.00	1.10	6.20	0.07

Note. SBSK-P = parental secure base script knowledge; SBSK-C = child secure base script knowledge; CBQ = child behavior questionnaire, discomfort subscale; SD = standard deviation.

Figure 7.1

Association between parental sensitivity and children's secure base script knowledge moderated by child discomfort



Note. For visual purposes, the discomfort scale was dichotomized into high discomfort (1 standard deviation or more above the mean) and low discomfort (1 standard deviation or more below the mean).

Table 7.2

Bivariate Pearson correlations between the study variables

Variable	1	2	3	4	5	6	7	8
1. Age child								
2. Sex child	.01							
	[-.09, .12]							
3. Gender parent	.07	-.02						
	[-.03, .17]	[-.11, .07]						
4. SBSK-P	.06	-.02	.04					
	[-.05, .17]	[-.11, .08]	[-.06, .14]					
5. Sensitivity	.05	-.00	-.08	**.17**				
	[-.05, .16]	[-.10, .10]	[-.18, .01]	**[.07, .27]**				
6. Discipline	-.04	.03	**-.11**	.05	**.28**			
	[-.15, .06]	[-.06, .13]	**[-.20, -.01]**	[-.06, .15]	**[.19, .37]**			
7. SBSK-C	.00	**.28**	-.05	.01	.07	-.01		
	[-.11, .11]	**[.18, .38]**	[-.16, .07]	[-.11, .12]	[-.04, .18]	[-.12, .10]		
8. PGS	.04	.07	.08	-.06	.02	.04	.09	
	[-.08, .16]	[-.05, .18]	[-.04, .19]	[-.18, .06]	[-.10, .13]	[-.08, .15]	[-.04, .21]	
9. Discomfort	-.03	**.20**	.01	**-.12**	.02	.03	.02	.08
	[-.13, .08]	**[.10, .29]**	[-.09, .11]	**[-.22, -.02]**	[-.08, .12]	[-.08, .13]	[-.10, .14]	[-.04, .20]

Note. Bold estimates are significant at $p < .05$; SBSK-P = parental secure base script knowledge; SBSK-C = child secure base script knowledge; CBQ = child behavior questionnaire, discomfort subscale; PGS = Polygenic score of Differential Susceptibility.

was not significant (β = .01, t(473) = 0.49, p = .625, based on ± 1 *SD*).

Sensitivity analyses

In sensitivity analyses, we repeated the main analyses using only participants with complete data, with similar results (see Table S7.5).

Discussion

In this study, we investigated the role of differential susceptibility in the transmission of secure base script knowledge from parents to their children. In line with the theoretical framework of intergenerational transmission of attachment mediated by parental sensitivity, we found an association between parental sensitivity and children's attachment representations. Children who experienced higher levels of parental sensitivity had more secure base script knowledge. This association was moderated by temperamental discomfort. The moderation analysis revealed that children with higher scores on discomfort (as an index for higher susceptibility) had less SBSK when their parents exhibited lower levels of sensitivity, but not more SBSK when their parents demonstrated higher levels of parenting sensitivity. This is support for a diathesis-stress interpretation rather than support for the differential susceptibility model, as we found a robust interaction between parental sensitivity and child discomfort, but children scoring higher compared to lower on discomfort do not differ from each other in SBSK at the higher end of parental sensitivity. The diathesis-stress model posits vulnerability or risk factors which moderate the association of parenting with child development in a cumulative-risk manner. The combination of risk factors, in our case temperamental discomfort and less sensitive parenting, predicted less secure attachment as measured with the ASA. In contrast to the differential susceptibility hypothesis, the diathesis-stress model assumes that difficult temperament might act "for worse" in a less optimal environment but not "for better" in a more optimal environment. We should note, however, a restriction of range at the high end of the SBSK scores (our highest individual child SBSK score based on the mean of the three stories was 4.67 on a scale from 1 to 7), maybe because the children were relatively young. Our sample was on average 1.5 to 2 years younger compared to previous studies with the ASA in middle childhood (the average age

in our sample was 9.09 years, compared to 10.91 in Waters et al. (2022) and 10.60 and 10.93 in Bosmans et al. (2014)). We noticed that some children experienced difficulties telling a story based on the word prompts. The task might thus have been too cognitively challenging for some children and consequently interfering with their ability to incorporate attachment scripts into their narratives (Bosmans et al., 2014; Waters et al., 2022). Visual inspection showed that there might be a "for better" susceptibility effect that we failed to capture because of this restriction in range of SBSK scores. It should be noted that when we explored the role of parental sensitive discipline, we failed to discover a pattern indicative of diathesis-stress or differential susceptibility.

Intriguingly, our findings deviate from what we expected based on a recent Individual Participant Data meta-analysis (Verhage et al., 2018) and a conventional meta-analysis (Verhage et al., 2016), as we did not observe a direct link between parental and child attachment and equivalence testing indicated that our estimate was equivalent to zero. Verhage et al. (2018) found a correlation of .29 between parental attachment representation (as measured with the AAI) and child attachment behavior (as observed in the SSP) in their Individual Participant Data meta-analysis and a correlation of .31 in the traditional meta-analysis. Although the discrepancy may seem surprising, several explanations for our findings can be offered. First, Verhage et al. (2016, 2018) explicitly excluded parental attachment representations measured with the ASA (Waters & Waters, 2006) because the ASA provides a continuous rating of security instead of a classification as secure versus insecure. Second, only dyads with children in infancy or early childhood were included in the meta-analyses (mean age 21 months; maximum age 6 years) whereas our sample was comprised of 9-year-olds. As children grow older, they get more diverse opportunities to form attachment bonds with other individuals than their parents, such as grandparents, childcare workers, teachers, or even parents of close friends. In middle childhood children acquire more cognitive capacities which facilitates the organization of multiple attachment figures in their representation of a growing number of attachments (Bakermans-Kranenburg, 2021). The expanding attachment network might decrease the influence of parental attachment representations (Dagan et al., 2021).

Third, in addition to the age difference in the children assessed, it is worth considering the temporal gap between the attachment assessment of the parent and the subsequent assessment of the child. In our longitudinal study, a time span of four years elapsed between the parental assessment and the child assessment of SBSK. The mean age of children in Verhage et al. (2018) was 21 months and the AAI was conducted either in the third trimester of the pregnancy or postnatally, which means that the time interval in the included studies was always less than 2 years, with a number of studies having assessed parent and child attachment concurrently. When longer time intervals are associated with lower convergence of parent and child attachment, the transmission of attachment (representations) might be less stable over time than previously assumed. This would be in line with longitudinal investigations revealing at most moderate stability of attachment security over an individual's childhood (Fraley, 2002; Groh et al., 2014; Pinquart et al., 2013). Additionally, the notion that attachment consists of a trait-like and a state-like component has been proposed (Bosmans et al., 2014; Gillath et al., 2009). The state-like component pertains to children's expectations, emotions and thoughts in the present moment. A recent conflict with a parent may affect the state-like component of attachment. This might have played a role, where half of the children were observed in the Do-Don't task before their SBSK was measured.

Our study has some limitations. We used three attachment stories in the ASA instead of four to minimize participant burden, which might have affected the reliability of the measure in adults. However, three stories are usually used in the child version of the ASA, and our ICCs and Cronbach's alpha were acceptable, which is why we think that omitting one attachment story might not have impacted the validity of this measure (Waters et al., 2015; Waters et al., 2019). Second, we should note that we assessed only one parent's attachment representations. Although this was the primary parent who reported to spend the most time with the children, the other parent might also play a crucial role in the transmission of attachment (representations) from parent to child. Future research should explore the transmission of attachment in the family context and include all major caregivers in their study (as in Dagan et al., 2021). Lastly, in the pre-registration of this study, we specified the analyses we planned to perform, but due to non-convergence and/or

inadequate model fit we had to revise our analytic strategy. Note that all deviations have been reported in a time-stamped addendum to the pre-registration. Given these limitations, it is crucial to replicate our study in a somewhat older age group.

Conclusion

In our longitudinal twin study with observed parenting and attachment measures, we did not find direct transmission of attachment operationalized as SBSK. However, we found parental sensitivity predicted child secure base script knowledge, and that temperamental difficulty, measured by the discomfort subscale of the CBQ moderated the association between parental sensitivity and attachment representations. On the premise of replication, we may conclude that temperament plays a role in explaining the variability in the intergenerational transmission of secure base script knowledge, in line with a diathesis-stress pattern instead of a differential susceptibility model.

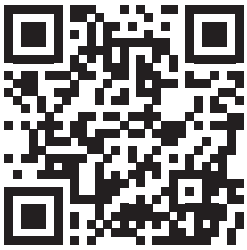
Acknowledgements

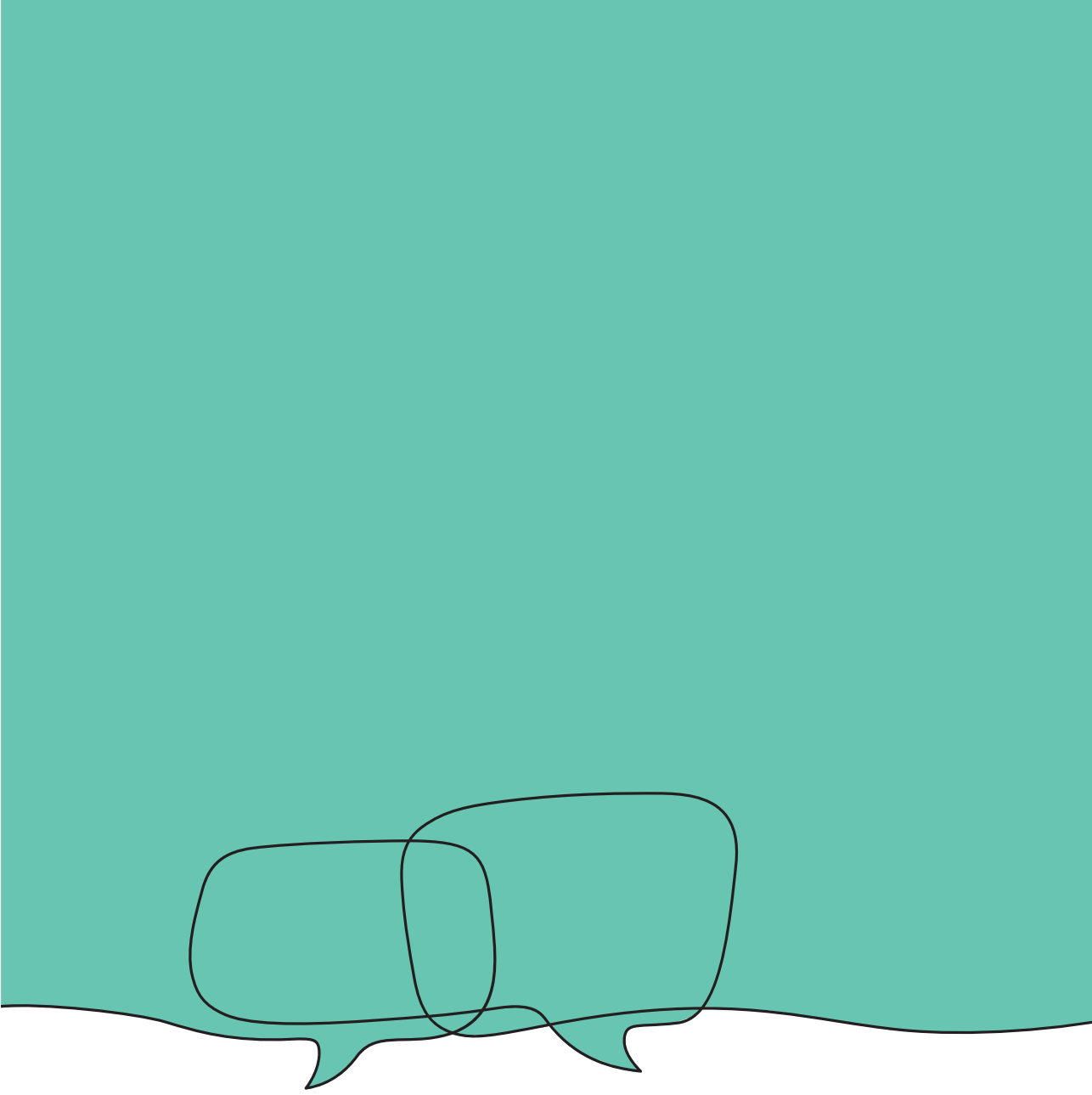
We thank the participating families for their enthusiastic involvement in the Leiden Consortium on Individual Development (L-CID). We are also grateful to the data-collection and data-processing team, including all current and former students, research assistants, PhD students and post-doctoral researchers for their dedicated and invaluable contributions. Marinus van IJzendoorn, Eveline Crone and Marian Bakermans-Kranenburg designed the L-CID experimental cohort-sequential twin study “Samen Uniek” as part of the Consortium on Individual Development (CID; Gravitation Grant 2013-2023 awarded by the Dutch Ministry of Education, Culture, & Science, and the Netherlands Organization for Scientific Research, NWO Grant Number 024.001.003). We thank Saskia Euser for training observers to code sensitive caregiving, discipline, and SBSK.

Supplementary Materials

All supplementary materials can be retrieved from:

http://tinyurl.com/Chapter7Supplement





GENERAL DISCUSSION

CHAPTER EIGHT

General discussion

Summary

In this thesis, I concentrated on determinants of parents' and children's response biases. Moreover, I focused on determinants of parenting behavior: parents' genes, their children's genes, parents' secure base scripts, and a randomized controlled trial (RCT) testing the effectiveness of a parenting intervention. I examined the determinants of children's attachment and their psychobiology, focusing on children's genes, their parents' behavior and effects of the parenting intervention. The studies complement each other by highlighting different aspects of parent-child relationship and children's psychobiological development.

First, in Chapter 2, I examined the role of genes in response bias of parents and their twin children. Then, I investigated whether genetic predispositions to socio-economic and cognitive factors (i.e., educational attainment, IQ and income) can predict sensitive parenting in mothers of the GenR study (Chapter 3). In Chapter 4, I focused on the heritability of sleep and cortisol in children. Then, in Chapter 5, I examined whether a brief attachment-based video-feedback intervention (VIPP-SD) could increase parental sensitivity and sensitive discipline in parents of school-aged twin children. Chapter 6 expanded the research about VIPP-SD by investigating its effect on children's hair cortisol levels and conduct problems. In Chapter 7, I examined whether the transmission of attachment representation is mediated by parental sensitivity and sensitive discipline and whether children are differentially susceptible to parental sensitivity and sensitive discipline. In this general discussion, I aim to provide global interpretations of the findings across studies and I place the specific findings of each of the studies in the larger context of parenting research and child development.

Reliability of self- or parent-report questionnaires

In psychological research, questionnaires have played a crucial role for years. They offer a convenient and cost-effective way of collecting data. However, their reliability has been questioned, especially when comparing self-reports to direct observations (Prince et al., 2008). Using a twin design, we identified that a significant portion of the variance in the acquiescence bias among children can be

attributed to genetic factors. Identical twin pairs exhibited more similar responses than non-identical twins. Additionally, we observed a notable effect of molecular genetic predictors on the acquiescence response bias. We examined the role of three highly correlated polygenic scores (i.e., PGSs for educational attainment, IQ, and income) in the response bias of parents and their twin children. A higher score on the latent factor of the polygenic scores of educational attainment, iQ, and income was associated with less bias in agreeing to impossible statements.

These findings align with prior research on response biases in questionnaire studies. Even more so, we uncovered evidence for genetic determinants of individual differences in the acquiescence bias, using behavior genetics as well as molecular genetic methods. Respondents are genetically predisposed to be more (or less) biased in their answers to absurd questions. This genetic predisposition poses a greater challenge than just a tendency to answer in a socially acceptable way. The genetic core of the bias cannot be influenced by exposure or outcome in developmental studies. Therefore, the bias will be a classic confounder inflating the associations between exposures and outcomes.

Genetic factors and gene-environment correlations in parenting

As a successor of candidate gene studies, polygenic scores have been a promising innovation in the field of parenting and child development research. In our study, we replicated the finding of Wertz et al. (2019) that observed sensitive parenting was significantly predicted by a polygenic score of educational attainment. However, we extended the study by including three highly correlated polygenic scores in the cognitive domain, namely the PGSs for educational attainment, IQ, and income. The advantage of such a higher-order aggregate might be better reliability and broader (ecological) validity.

Another important and new aspect of our study is that differences in observed sensitivity were explained, in part, by genetic differences between children. The inclusion of child genetic effects in the model increased the explained variance in observed sensitivity, highlighting an important path between child genotype and parental behavior. This finding is in line with Belsky's (1984) process model of parenting as well as Sameroff's (2009) transactional model of development,

wherein child characteristics play a role in parent-child interactions. Such evocative genetic effects have important implications. Parental sensitivity has often been considered to be a parental trait more or less independent of child factors (Cooke et al., 2022; Steele et al., 1996). Our findings suggest otherwise: perhaps moderately sensitive parents respond differently to their children depending on, for example, the temperament of the child: such parents may find it easier to respond sensitively to easy-going children compared to irritable children. The sensitive interactions of parents with their children become essentially dyadic with exchange of information, signals and emotions where parents are the responsible agents but children are active participants. These results may inform personalized parenting advice based on child characteristics.

We also included relevant maternal phenotypes (i.e., maternal education, household income, and maternal IQ) as predictors of maternal sensitivity in toddlerhood and early childhood. At both ages we saw evidence of positive associations, with higher income and maternal IQ predicting more observed sensitive parenting. The addition of maternal PGS-EA and PGS-EDINQ increased explained variance over and above the related phenotypes, emphasizing the role of PGSs as a valuable tool in family studies. PGSs may reflect a genetic proclivity of these phenotypes which does not necessarily mean that these have (already) been realized in real-life. This is important to consider when interpreting results of polygenic score studies, wherein genetic proclivities do not always and inevitably equal their phenotypic counterparts, especially when it comes to complex behavioral phenotypes.

In sum, our study shows the added value of genetic data in developmental psychology for the following reasons. First, it is possible to measure genetic sensitivities which are not (yet) translated into behavior and therefore impossible to measure using questionnaires, experiments or observations. Insights from genetically informed studies might guide us in prevention of unfavorable behavioral outcomes. Second, the prediction of behavior by polygenic scores over and above behavioral phenotypes is a reminder that not all facets of the specific phenotype are measured. Third, parenting and child development are complex phenotypes which are at the same time important factors in everyone's lives. Including genes in our research gives us the possibility to disentangle under which (genetic or

environmental) circumstances specific behaviors occur. Subsequently, parenting interventions can focus on improving the environments or preventing behaviors which are most malleable.

Intergenerational transmission of attachment

The term "intergenerational transmission of attachment" has been used to describe the transmission of parental attachment representations to their children's attachment relationship. Previous research has suggested parental sensitivity as a mediator in this relationship (Verhage et al., 2016). We examined this issue in our longitudinal twin study. As expected, we found that parental sensitivity predicted children's attachment representations. Children with more sensitive parents had more secure attachment representations. In contrast with Verhage et al.'s (2016) Individual Participant Data meta-analysis, we did not discover a direct transmission of parental attachment representations to children's attachment representations.

One difference between the Individual Participant Data meta-analysis and our study were the measures used. In Verhage et al. (2016) the transmission of attachment was investigated between the Adult Attachment Interview (AAI) and the Strange Situation Procedure (SSP) whereas we investigated the transmission using the Attachment Script Assessment in both parents and the children. The AAI has been found to significantly correlate with secure base script knowledge with correlation coefficients ranging from .26 to .62. (Coppola et al., 2006; Dykas et al., 2006; H. S. Waters & Rodrigues-Doolabh, 2001). This indicates that although there is overlap in the measures, it is also clear that both measures do not exactly measure the same underlying construct. In children, there was no significant correlation between the SSP and secure base script knowledge in a study of internationally adopted Chinese girls (Finet et al., 2020), suggesting that there might be differences between attachment security in infancy and attachment representations in childhood. However, this is difficult to investigate as neither the AAI nor the SSP are suitable measures for the attachment (representations) of children.

Nevertheless, these differences might explain the lack of direct transmission. Another difference with Verhage et al.'s (2016) study was that instead of investigating the transmission in infancy, we investigated it in a sample of 9-year-olds. In this

developmental period, children might have acquired sufficient cognitive capabilities to form attachment representations of multiple attachment figures which might consequently lead to a decrease of the influence of an individual attachment figure (Bakermans-Kranenburg, 2021; Dagan et al., 2021). 9-year-olds have most likely encountered multiple attachment figures, such as the other parent (mostly the father in our sample), grandparents, as well as caregivers in a child care setting and teachers in primary schools. Therefore, children might draw upon the combination of attachment experiences and the resulting secure base script knowledge might not necessarily depict secure base script knowledge based on experiences with the caregiver in question.

According to Bakermans-Kranenburg (2021), the size of the attachment network (and therefore the members of this network) to which children are able to attach to is limited by the cognitive capabilities of the child. The ongoing development of cognitive abilities during childhood might increase the size of this attachment network. Older children should have better cognitive abilities to incorporate the behavior and characteristics of attachment figures in their environment into attachment representations. At the same time, these attachment experiences with several caregivers might become integrated into an overarching attachment representation. In this case, the secure base script knowledge would not solely represent the attachment representation to the specific caregiver we included in our study. Supporting evidence for this idea comes from Waters and colleagues (2015) who investigated whether secure base script knowledge of adolescents is generalized across a variety of attachment relationships and found an overarching higher dimension factor of secure base script knowledge across several attachment relationships. When and how such a generalized view of secure base script knowledge develops, is an important question which has, to our knowledge, not yet been explored.

Interventions to improve parenting

We conducted an RCT using the Video-feedback to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer et al., 2017). We did not find a significant effect of the intervention on parental sensitivity or sensitive discipline. Besides

parental sensitive behavior, we examined whether VIPP-SD impacted parents' attitudes towards sensitivity and sensitive discipline. Parents' attitudes towards sensitivity were significantly improved by VIPP-SD but not parents' attitudes towards sensitive discipline. The lack of a significant intervention effect on observed parental sensitivity and sensitive discipline was surprising given previous meta-analytic findings on the effectiveness of the VIPP-SD program (Juffer et al., 2017). However, VIPP-SD was originally designed for children up to three years of age and we adapted the intervention for school-aged children. One possible explanation for the lack of effect would be that VIPP-SD is indeed not effective in parents of school-aged children. These parents have had many years of shaping their own parenting behavior and it may be harder to change behavior that developed over such a long time compared to behavior of parents with younger children. One might therefore conclude that the effectiveness of VIPP-SD is limited to infancy and early childhood, at least in upper- and middle-class families with twins.

However, another explanation for the lack of significant effect on behavior lies in the measurement of parental sensitivity and sensitive discipline. In the early childhood cohort of the L-CID study, VIPP-SD has been found to be effective in promoting sensitive discipline (Euser et al., 2021), and in this study, the Don't touch task was used. In this task, the child was not allowed to touch a range of toys and parents were free to focus on their child and preventing the child from touching the toys. In contrast, we used a combined Do/Don't task in our study. In this task parents had to ensure that their child did a boring task and at the same time keep their child from watching an attractive video. Thus, the parents had to 1) make sure that the child stayed engaged in the "Do" task, 2) make sure the child did not (secretly) watch the video, and 3) watch the video themselves. The situation might have been too complicated to utilize their newly learned parenting skills.

Alternatively, VIPP-SD might be effective in improving other parenting constructs apart from sensitivity and sensitive discipline in parents of school-aged children. If, at the same time, effects of the intervention on children are found, this implies that the mechanism of improving child behavior is not to be found in parental sensitivity and sensitive discipline. Other, unmeasured parenting dimensions may have been affected which seems to be likely considering the effects of VIPP-SD on

child behavior and physiology in the current study. Previous RCTs which included other parenting constructs might provide some inspiration for future research: the VIPP, adapted for different at-risk groups, increased self-efficacy (Platje et al., 2018; Poslawsky et al., 2015), decreased harsh discipline (Pereira et al., 2014), improved familial relational functioning (Negrão et al., 2014) and decreased parenting stress (Hodes et al., 2017). However, the question about which parenting dimension in normative samples (i.e., parents and children not at-risk) might be influenced by VIPP-SD, is still open for investigation.

Genetic factors in child psychobiology

We employed twin models to investigate the heritability of sleep and cortisol. We found that sleep duration, sleep efficiency and wake episodes were moderately heritable. A high genetic correlation was found between sleep duration and sleep efficiency, and also between sleep efficiency and wake episodes which is in line with previous research (Sletten et al., 2013). We also found high genetic correlations between several aspects of sleep, i.e., between sleep duration and sleep efficiency and between sleep efficiency and wake episodes, but not between sleep duration and wake episodes. In line with recent genome-wide association studies (Dashti et al., 2019; Doherty et al., 2018; Lane et al., 2017), our findings indicate that distinct genetic factors might be implicated in sleep duration compared to waking episodes. At the same time these genetic factors might as well be correlated, therefore, more twin studies and genome-sequencing research are needed to shed light on the contribution of different genes to the variation in sleep duration, sleep efficiency and wake episodes.

Regarding cortisol, we found no significant heritability of morning cortisol levels. The twin literature shows a large variability in heritability estimates, depending on age, time of sampling, and type of measurement (urine, blood, or saliva) with heritability estimates ranging from 0 to 88% (Bartels et al., 2003). However, Neumann and colleagues (2016) examined the SNP heritability (i.e., the variance explained by common autosomal single nucleotide polymorphisms together) of plasma and salivary cortisol in several studies. They found no significant SNP heritability for cortisol in any of the used samples. Cortisol levels seem to fluctuate on a day-to-day

level depending on several situation-specific environmental influences (Ross et al., 2014; Shirtcliff et al., 2012). These studies highlight the possibility of parents to co-regulate cortisol levels of their children which we investigated using the VIPP-SD program.

Child psychobiology

Although we did not find significant effects of VIPP-SD on observed sensitivity and sensitive discipline in the L-CID middle childhood cohort, the intervention was successful in decreasing children's conduct problems and hair cortisol levels in our population-based sample of school-aged twin children. This strengthens our speculations that the measures used to assess parenting behaviors were inadequate or did not capture parenting constructs which were improved by VIPP-SD. Our findings are in line with Van IJzendoorn's (2022) recent meta-analysis, even though he did not find a significant effect of VIPP-SD on externalizing behaviors. However, our effect size was comparable to the combined effect size of the meta-analysis. One interesting difference between our study and the studies included in the meta-analysis is the age of the children included in the studies. In the majority of the included trials in the meta-analysis, children were infants or toddlers before the age of four whereas in our study, children were above age four. Possibly, VIPP-SD specifically influences parenting behaviors that are adaptive in decreasing externalizing behaviors typically shown in childhood rather than infancy although this question needs to be investigated in future research. Based on previous research, it seems like VIPP-SD has effects on parents' behavior when their children are younger (i.e., infants or toddlers) and effects on children's behavior but not their parents' behavior when children are in early or middle childhood. In future research, we need to examine whether these puzzling findings can be explained by the use of different parenting constructs as well as which parenting constructs are most affected and influential.

Not only did the intervention decrease conduct problems, but in the intervention group also children's hair cortisol concentrations decreased after the intervention compared to the control group. Our findings are congruent with Poehlmann-Tynan et al. (2020) who found that after a cognition-based compassion training (CBCT) with 25 parents and their children aged between 4 months and 5

years, child hair cortisol levels were significantly lower than those of children in the waitlist control group. Here we replicated these results in a well-powered twin study, with an attachment-based parenting intervention. Our findings are in contrast with the overall results of a recent meta-analysis (Martins et al., 2020), which concluded that parenting intervention did not exert an effect on children's cortisol levels. However, the majority of these studies used salivary cortisol measures, such as the cortisol awakening response or cortisol reactivity in response to a stressor. These measures have been found to only weakly correlate with hair cortisol and be subject to a variety of confounding factors, such as food intake or physical activity (Neumann et al., 2016). Moreover, high salivary cortisol levels in a stressful situation constitute an adaptive feature of the human body, with only long-term chronic cortisol levels being associated with adverse health outcomes (McEwen, 2005; Miller et al., 2007). Our results support the idea that hair cortisol levels seem to be a better option when investigating the effects of parenting interventions on chronic cortisol levels of children.

Differential susceptibility

In this thesis, we investigated the differential susceptibility theory in three of the studies. In the randomized controlled trial, we examined whether parents were differentially susceptible to the intervention effects depending on their temperamental reactivity. We did not find support for the differential susceptibility hypothesis. Focusing on child effects, we tested differential susceptibility to the positive effects of the intervention using two polygenic scores. We did not find support for differential susceptibility to parenting with respect to conduct problems or hair cortisol concentrations. In the third study, we investigated the role of differential susceptibility measured by children's temperament in the transmission of attachment representations from parents to their children. We did not find support for the differential susceptibility hypothesis, but the findings were in line with the diathesis-stress model.

One possible explanation for the first two findings is domain-specificity of differential susceptibility. We speculate that the parent and child temperamental measures of differential susceptibility markers, as well as the two PGSs in our study

might be indicators of differential susceptibility, but specific to the outcome domains for which they were developed, and not general or for the specific exposure or outcome domain of our study. The polygenic score of differential susceptibility that we used was developed by Keers and colleagues (2016) and is a polygenic score based on GWAS data for anxiogenic environments. In Keers et al. (2016) this polygenic score moderated the association between child-reported parenting and children's internalizing emotional problems. Using the same polygenic score, Lemery-Chalfant et al. (2018) found that the PGS moderated the association between effects of the Family Check-up intervention and internalizing psychopathology. Considering that we did not find moderation of the effect of VIPP-SD on conduct problems, we may conclude that this polygenic score seems more specific to internalizing behavior. We had speculated that susceptibility towards effects of parenting on internalizing and externalizing behaviors might be indicated by one underlying PGS, but we were not able to corroborate this assumption. Further research is needed to investigate whether this PGS is a susceptibility marker for any outcome outside the internalizing domain.

Furthermore, we explored the polygenic score (PGS-SESA) which was based on the GWAS of "sensitivity to environmental stress and adversity" cluster of neuroticism (Nagel et al., 2020) as a possible differential susceptibility marker. However, in our study, the polygenic score did not emerge as a differential susceptibility marker. This polygenic score might be specific to the cognitive perception of stress and adversity rather than neurobiological measures of stress. More research on the PGS-SESA is clearly required to draw evidence-based conclusions about its role as marker of differential susceptibility to different environments and with various outcomes.

In the third study, we did not find a pattern compatible with differential susceptibility when we investigated the relationship between parental sensitive discipline and children's attachment representations which strengthens the assumption of domain-specificity in differential susceptibility. However, our findings regarding parental sensitivity and children's attachment representations were in line with the diathesis-stress model instead of the differential susceptibility hypothesis. We found a robust interaction between parental sensitivity and child discomfort, wherein children scoring higher compared to lower on temperamental discomfort

did have less secure attachment representations when parental sensitivity was low. However, children scoring higher compared to lower on discomfort did not differ from each other in their attachment representations at the higher end of parental sensitivity. According to the diathesis-stress model, vulnerability factors moderate the association between parenting and child development in a cumulative-risk manner. In this study, the combination of less sensitive parenting and higher temperamental discomfort predicted less secure attachment representations. In the diathesis-stress model, there is only a "for worse" effect and not a "for better" effect, in contrast with the differential susceptibility theory. We made use of the discomfort subscale of the temperament in middle childhood questionnaire (Simonds, 2006). This scale measures negative affect with regard to sensory stimulation, for example intensity of light, sound or texture. A meta-analysis of Slagt et al. (2016) investigated different susceptibility markers to parenting used in previous research and found that difficult temperament and negative emotionality in infancy, but not surgency or effortful control, emerged as susceptibility markers, indicating that many temperamental dimensions are not useful in investigating differential susceptibility in the parenting domain. Unfortunately, this meta-analysis did not include temperamental discomfort although discomfort is part of the dimension of negative emotionality, among fear, worry, sadness, anger, frustration, and irritability (Rothbart, 2004). Possibly, temperamental discomfort is a suitable marker for differential susceptibility in infancy only, as negative emotionality seems to be.

Before drawing definite conclusions, we need to consider a methodological aspect of the current study: the range of scores on the attachment representation measure was limited at the high end of the scores. Perhaps the children in our sample were too young to be able to demonstrate high levels of secure base script knowledge. Compared to other studies which made use of the same measure and reported a full range of scores, our mean age was lower as well as the upper end of the age range (Bosmans et al., 2014; T. E. A. Waters et al., 2022). Visual inspection suggests that there might be a differential susceptibility effect that we were unable to capture because of the restriction in the range of our attachment representation scores.

Limitations

Some limitations should be noted with regard to this thesis. One of the major drawbacks in this thesis is the use of samples with predominantly mother-child dyads as participants instead of including fathers as well, or, in the best case, both caregivers within a family. Fathers have been underrepresented in parenting research for a long time, and even though their roles in the context of parenting seem to be changing in society, contemporary cohort studies still rely mostly on mothers. Besides research on fathering, we also lack knowledge on the interactions between parents and their children as a trio (but see Witte et al. (2020, 2021) for research on trios). It is crucial to investigate whether there are buffering effects when one parent is more sensitive and the other is less sensitive or whether one plus one equals three, i.e., whether interaction effects of parents' individual parenting behaviors amplify the effect on their children.

Another limitation to be discussed is the potential weakness of polygenic scores. Although the rise of GWAS had many positive effects on the progress of research, the use of polygenic scores based on GWAS is not flawless. The quality of a polygenic score depends on several aspects: the availability of a GWAS, the sample size (and related, power) of that GWAS, as well as the quality of the phenotype measures used for the GWAS. Unfortunately, no GWAS on parenting or differential susceptibility have been conducted to date, which is why we made use of proxy PGSs in the quest for the genetic predisposition of parenting and differential susceptibility. As a consequence, the variance explained by the polygenic score is very low and one may question the value of the results if a PGS explains around one or two percent of the variance in, for example, parental sensitivity. Moreover, many GWASs are based on self-report questionnaires, sometimes even single questions. In behavioral research, and specifically parenting research, there have been criticisms about the use of questionnaires as these are often unreliable due to, for example, social desirability biases (Kagan, 2007; Runze & Van IJzendoorn, 2023). As long as GWASs of complex phenotypes are based on (single-item) questionnaires, these criticisms apply to GWAS and resulting polygenic scores as well.

Future directions

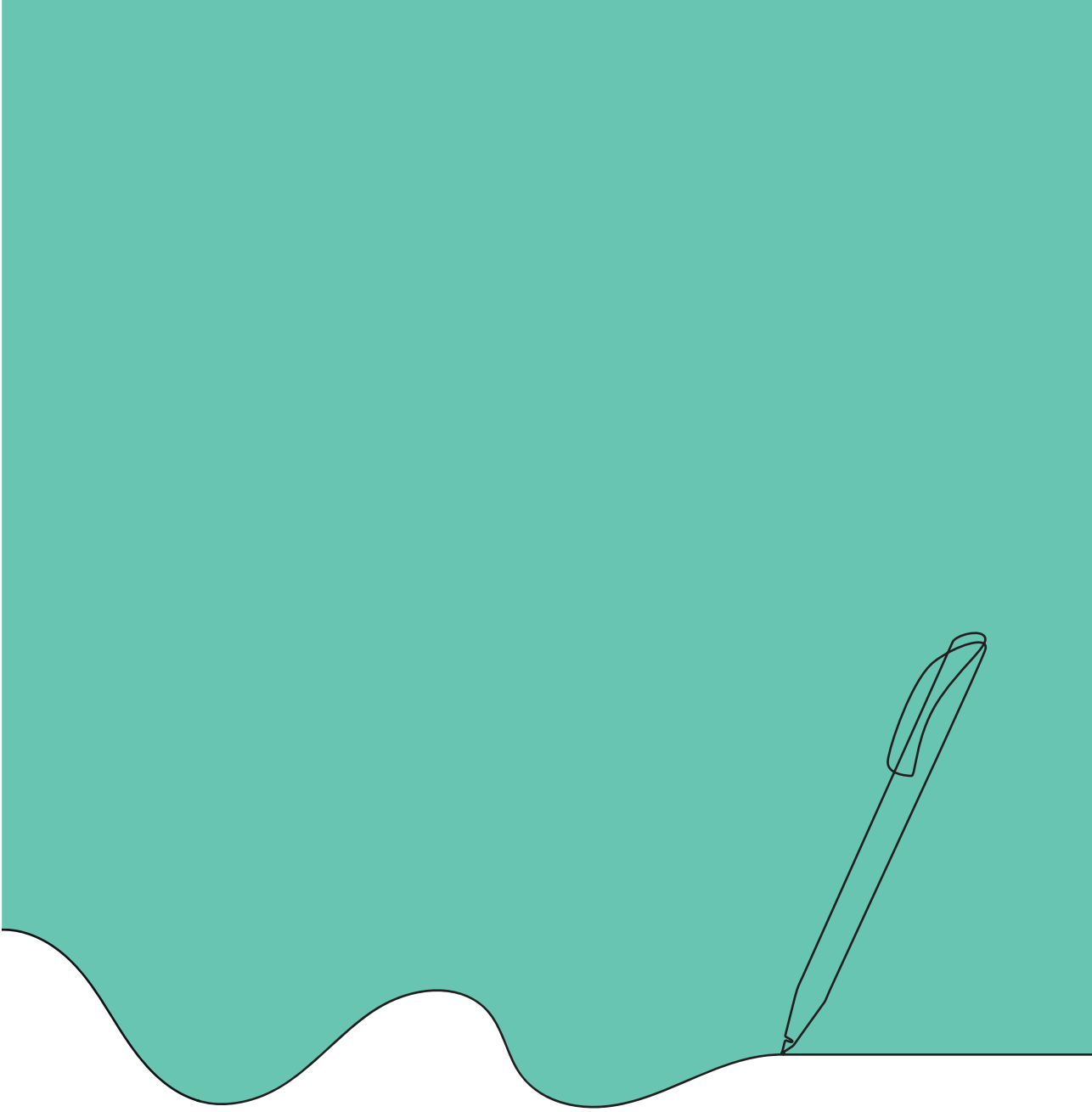
More progress is needed in GWAS research on observed parenting behaviors and differential susceptibility to facilitate the investigation of causality in parenting research. To examine causal relationships, a GWAS of parenting is a crucial next step to be taken to expand current gene-environment parenting research enabling Mendelian Randomization (MR) studies. Mendelian Randomization is a statistical technique to investigate causal relationships between a set of variables. MR makes use of measured genetic variants related to, for example, parental sensitivity to assess the causal effect of sensitivity on, for example, child behavioral problems. A polygenic score of parenting could be used as a genetic instrument to enable MR in developmental and parenting research.

An avenue for future research regarding differential susceptibility is the hypothesized domain-specificity of differential susceptibility (Belsky et al., 2022; Zhang et al., 2023). From an evolutionary standpoint, we might expect a certain specificity of differential susceptibility. Humans' responsiveness may have evolved to different (types of) environmental cues depending on their reproductive and survival needs in line with conditional adaptation (Boyce & Ellis, 2005). There may have been distinct challenges in different domains of life which have given rise to the emergence of distinct susceptibility markers. This specificity can occur at several levels: At the level of the exposure and at the level of the outcome. For example, a "susceptible" child might be susceptible to their parents' behavior but not susceptible to their peers' behavior towards them. On a more fine-grained level, it might even be the case that a child is susceptible to their parents' sensitive responsiveness but not to their parents' discipline. The same might be true when it comes to outcomes. A child might be susceptible to parental behaviors in affecting the occurrence of externalizing behaviors but not internalizing behaviors or vice versa. A first couple of studies do provide support for domain-specific susceptibility but more research is needed to disentangle which susceptibility markers capture susceptibility in which exposure and outcome domain (Belsky et al., 2022; Sayler et al., 2022; Zhang et al., 2023). Current knowledge about specificity is too scarce to draw definite conclusions when susceptibility markers do not seem to play a moderating role. Moreover, currently available PGS of differential susceptibility have failed to reliably detect

differential susceptibility effects. However, PGS have the potential to play a crucial role in the advancement of differential susceptibility research. GWAS of domain-specific differential susceptibility factors are needed to inform genetic differential susceptibility models.

Conclusions

In this thesis, I focused on determinants of parenting behavior: parents' genes, their children's genes, parents' secure base scripts, and a randomized controlled trial (RCT) testing the effectiveness of a parenting intervention. I investigated determinants of children's attachment and their psychobiology, focusing on children's genes, their parents' behavior and effects of the parenting intervention. We found that maternal sensitivity is predicted by mother and child genetic predispositions to socio-economic and cognitive factors (i.e., educational attainment, IQ, and income). In children, sleep was found to be moderately heritable whereas variability in cortisol levels was mostly explained by environmental factors. Although the brief attachment-based video-feedback intervention (VIPP-SD) did not increase parental sensitivity and sensitive discipline in parents of school-aged twin children, we found effects of VIPP-SD on the child level: families who received the VIPP-SD had children with lower conduct problems and hair cortisol levels. We found no transmission of attachment representation from parents to their children, but we did find that parental sensitivity predicted child attachment representations as measured with the ASA. Future research can make significant steps in parenting research by incorporating genome-wide association studies of observed parenting measures and making use of polygenic scores in the quest for domain-specific differential susceptibility.



APPENDICES

CHAPTER NINE

0-9

23andMe Research Team, COGENT (Cognitive Genomics Consortium), Social Science Genetic Association Consortium, Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genetics, 50(8), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

A

Acebo, C., Sadeh, A., Seifer, R., Tzischinsky, O., Wolfson, A. R., Hafer, A., & Carskadon, M. A. (1999). Estimating Sleep Patterns with Activity Monitoring in Children and Adolescents: How Many Nights Are Necessary for Reliable Measures? Sleep, 22(1), 95–103. https://doi.org/10.1093/sleep/22.1.95

Achenbach, T. M., & Rescorla, L. A. (2000). Manual for the ASEBA preschool forms & profiles: An integrated system of multi-informant assessment. ASEBA.

Ainsworth, Bell, S. M., & Stayton, D. F. (1974). Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In M. P. M. Richards (Ed.), The integration of a child into a social world (pp. 99–135). Cambridge University Press.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (Eds.). (1978). Patterns of attachment: A psychological study of the strange situation (Vol. xviii). Lawrence Erlbaum.

Alexander, N., Osinsky, R., Schmitz, A., Mueller, E., Kuepper, Y., & Hennig, J. (2010). The BDNF Val66Met polymorphism affects HPA-axis reactivity to acute stress. Psychoneuroendocrinology, 35(6), 949–953. https://doi.org/10.1016/j.psyneuen.2009.12.008

Alink, L. R. A., Mesman, J., Van Zeijl, J., Stolk, M. N., Juffer, F., Koot, H. M., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2006). The Early Childhood Aggression Curve: Development of Physical Aggression in 10- to 50-Month-Old Children. Child Development, 77(4), 954–966. https://doi-org.vu-nl.idm.oclc.org/10.1111/j.1467-8624.2006.00912.x

Avinun, R., & Knafo, A. (2014). Parenting as a Reaction Evoked by Children's Genotype: A Meta-Analysis of Children-as-Twins Studies. Personality and Social Psychology Review, 18(1), 87–102. https://doi.org/10.1177/1088868313498308

B

Bakermans-Kranenburg, M. J. (2021). The limits of the attachment network. New Directions for Child and Adolescent Development, 2021(180), 117–124. https://doi.org/DOI: 10.1002/cad.20432

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2008). Oxytocin receptor (OXTR) and serotonin transporter (5-HTT) genes associated with observed parenting. Social Cognitive and Affective Neuroscience, 3(2), 128–134. https://doi.org/10.1093/scan/nsn004

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2014). A sociability gene? Meta-analysis of oxytocin receptor genotype effects in humans. Psychiatric Genetics, 24(2), 45–51. https://doi.org/10.1097/YPG.0b013e3283643684

Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2015). The hidden efficacy of interventions: Gene×environment experiments from a differential susceptibility perspective. Annual Review of Psychology, 66, 381–409. https://doi.org/10.1146/annurev-psych-010814-015407

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Mesman, J., Alink, L. R. A., & Juffer, F. (2008). Effects of an attachment-based intervention on daily cortisol moderated by dopamine receptor D4: A randomized control trial on 1- to 3-year-olds screened for externalizing behavior. Development and Psychopathology, 20(3), 805–820. https://doi.org/10.1017/S0954579408000382

Bartels, M., De Geus, E. J. C., Clemens, K., Sluyter, F., & Boomsma, D. I. (2003). Heritability of Daytime Cortisol Levels in Children. Behavior Genetics, 33(4), 421–433.

Bartels, M., Van den Berg, M., Sluyter, F., Boomsma, D. I., & de Geus, E. J. C. (2003). Heritability of cortisol levels: Review and simultaneous analysis of twin studies. Psychoneuroendocrinology, 28(2), 121–137. https://doi.org/10.1016/S0306-4530(02)00003-3

Bates, T. C., Neale, M. C., & Maes, H. H. (2019). umx: A library for Structural Equation and Twin Modelling in R. Twin Research and Human Genetics, 22, 27–41. https://doi.org/10.1017/thg.2019.2

Belsky, J. (1984). The Determinants of Parenting: A Process Model. Child Development, 55, 83–96.

Belsky, J. (2005). Differential susceptibility to rearing influence: An evolutionary hypothesis and some evidence. In B. J. Ellis & D. Bjorklund (Eds.), Origins of the social mind: Evolutionary psychology and child development (pp. 139–163). Guilford Press.

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, J., & Van IJzendoorn, M. H. (2017). Genetic differential susceptibility to the effects of parenting. Current Opinion in Psychology, 15, 125–130. https://doi.org/10.1016/j.copsyc.2017.02.021

Belsky, J., Zhang, X., & Sayler, K. (2022). Differential susceptibility 2.0: Are the same children affected by different experiences and exposures? Development and Psychopathology, 34, 1025–1033. https://doi.org/10.1017/S0954579420002205

Bentler, P. M. (1990). Comparative Fit Indexes in Structural Models. Psychological Bulletin, 107(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bernard, K., Hostinar, C. E., & Dozier, M. (2015). Intervention Effects on Diurnal Cortisol Rhythms of Child Protective Services–Referred Infants in Early Childhood: Preschool Follow-up Results of a Randomized Clinical Trial. JAMA Pediatrics, 169(2), 112. https://doi.org/10.1001/jamapediatrics.2014.2369

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T., Mehta, P., & Fox, J. (2011). OpenMx: An Open Source Extended Structural Equation Modeling Framework. Psychometrika, 76(2), 306–317. https://doi.org/10.1007/s11336-010-9200-6

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. Nature Reviews Genetics, 3(11), 872–882. https://doi.org/10.1038/nrg932

Booth, A. T., Macdonald, J. A., & Youssef, G. J. (2018). Contextual stress and maternal sensitivity: A meta-analytic review of stress associations with the Maternal Behavior Q-Sort in observational studies. Developmental Review, 48, 145–177. https://doi.org/10.1016/j.dr.2018.02.002

Bor, W., McGee, T. R., & Fagan, A. A. (2004). Early Risk Factors for Adolescent Antisocial Behaviour: An Australian Longitudinal Study. Australian & New Zealand Journal of Psychiatry, 38(5), 365–372. https://doi.org/10.1080/j.1440-1614.2004.01365.x

Borairi, S., Fearon, P., Madigan, S., Plamondon, A., & Jenkins, J. (2021). A mediation meta-analysis of the role of maternal responsivity in the association between socioeconomic risk and children's language. Child Development, 92(6), 2177–2193. https://doi.org/10.1111/cdev.13695

Borbély, A. A., & Achermann, P. (1999). Sleep Homeostasis and Models of Sleep Regulation. Journal of Biological Rhythms, 14(6), 559–568.

Bornstein, M., Putnick, D., & Esposito, G. (2017). Continuity and Stability in Development. Child Dev Perspect, 11(2), 113–119.

Bosmans, G., Bakermans-Kranenburg, M. J., Vervliet, B., Verhees, M. W. F. T., & van IJzendoorn, M. H. (2020). A learning theory of attachment: Unraveling the black box of attachment development. Neuroscience & Biobehavioral Reviews, 113, 287–298. https://doi.org/10.1016/j.neubiorev.2020.03.014

Bosmans, G., Van de Walle, M., Goossens, L., & Ceulemans, E. (2014). (In)variability of Attachment in Middle Childhood: Secure Base Script Evidence in Diary Data. Behaviour Change, 31(4), 225–242. https://doi.org/10.1017/bec.2014.18

Bosmans, G., Van Vlierberghe, L., Bakermans-Kranenburg, M. J., Kobak, R., Hermans, D., & van IJzendoorn, M. H. (2022). A Learning Theory Approach to Attachment Theory: Exploring Clinical Applications. Clinical Child and Family Psychology Review, 25(3), 591–612. https://doi.org/10.1007/s10567-021-00377-x

Bouchard Jr, T. J., & Propping, P. (Eds). (1993). Twins as a tool of behavioral genetics. John Wiley & Sons.

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. American Journal of Orthopsychiatry, 52(4), 664–678. https://doi.org/10.1111/j.1939-0025.1982.tb01456.x

Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. Development and Psychopathology, 17(02). https://doi.org/10.1017/S0954579405050145

Breitenstein, R. S., Doane, L. D., Clifford, S., & Lemery-Chalfant, K. (2018). Children's sleep and daytime functioning: Increasing heritability and environmental associations with sibling conflict. Social Development, 27(4), 967–983. https://doi.org/10.1111/sode.12302

Breitenstein, R. S., Doane, L. D., & Lemery-Chalfant, K. (2021). Children's objective sleep assessed with wrist-based accelerometers: Strong heritability of objective quantity and quality unique from parent-reported sleep. Sleep, 44(1), zsaa142. https://doi.org/10.1093/sleep/zsaa142

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., … Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Research, 47(D1), D1005–D1012. https://doi.org/10.1093/nar/gky1120

Buckley, T. M., & Schatzberg, A. F. (2005). On the Interactions of the Hypothalamic-Pituitary-Adrenal (HPA) Axis and Sleep: Normal HPA Axis Activity and Circadian Rhythm, Exemplary Sleep Disorders. The Journal of Clinical Endocrinology & Metabolism, 90(5), 3106–3114. https://doi.org/10.1210/jc.2004-1056

C

Cents, R. A. M., Kok, R., Tiemeier, H., Lucassen, N., Székely, E., Bakermans-Kranenburg, M. J., Hofman, A., Jaddoe, V. W. V., van IJzendoorn, M. H., Verhulst, F. C., & Lambregtse -van den Berg, M. P. (2014). Variations in maternal 5-HTTLPR affect observed sensitive parenting. Journal of Child Psychology and Psychiatry, 55(9), 1025–1032. https://doi.org/10.1111/jcpp.12205

Chang, L., Schwartz, D., Dodge, K. A., & McBride-Chang, C. (2003). Harsh Parenting in Relation to Child Emotion Regulation and Aggression. Journal of Family Psychology, 17(4), 598–606. https://doi.org/10.1037/0893-3200.17.4.598

Choi, S. W., Mak, T. S.-H., & O'Reilly, P. F. (2020). Tutorial: A guide to performing polygenic risk score analyses. Nature Protocols, 15(9), 2759–2772. https://doi.org/10.1038/s41596-020-0353-1

Choi, S. W., & O'Reilly, P. F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. GigaScience, 8(7), giz082. https://doi.org/10.1093/gigascience/giz082

Chrousos, G. P., & Kino, T. (2007). Glucocorticoid action networks and complex psychiatric and/or somatic disorders. Stress, 10(2), 213–219. https://doi.org/10.1080/10253890701292119

Cooke, J. E., Deneault, A.-A., Devereux, C., Eirich, R., Fearon, R. M. P., & Madigan, S. (2022). Parental sensitivity and child behavioral prolems: A meta-analytic review. Child Development, 93, 1231–1248. https://doi.org/10.1111/cdev.13764

Coppola, G., Vaughn, B. E., Cassibba, R., & Costantini, A. (2006). The attachment script representation procedure in an Italian sample: Associations with adult attachment Interview scales and with maternal sensitivity. Attachment & Human Development, 8(3), 209–219. https://doi.org/10.1080/14616730600856065

Costa, P. T., & McCrae, R. R. (1985). The NEO personality inventory: Manual, form S and form R. Psychological Assessment Resources, Inc.

Crandall, A., Deater-Deckard, K., & Riley, A. W. (2015). Maternal emotion and cognitive control capacities and parenting: A conceptual framework. Developmental Review, 36, 105–126. https://doi.org/10.1016/j.dr.2015.01.004

Crone, E. A., Achterberg, M., Dobbelaar, S., Euser, S., van den Bulk, B., der Meulen, M. van, van Drunen, L., Wierenga, L. M., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2020). Neural and behavioral signatures of social evaluation and adaptation in childhood and adolescence: The Leiden consortium on individual development (L-CID). Developmental Cognitive Neuroscience, 45, 100805. https://doi.org/10.1016/j.dcn.2020.100805

Culpin, I., Bornstein, M. H., Putnick, D. L., Sallis, H., Lee, R., Cordero, M., Rajyaguru, P., Kordas, K., Cadman, T., & Pearson, R. M. (2020). Specific domains of early parenting, their heritability and differential association with adolescent behavioural and emotional disorders and academic achievement. European Child & Adolescent Psychiatry, 29(10), 1401–1409. https://doi.org/10.1007/s00787-019-01449-8

D

Dagan, O., Sagi-Schwartz, A., & van IJzendoorn, M. H. (2021). Attachment networks to multiple caregivers: An introduction to a special issue. New Directions for Child and Adolescent Development, 2021(180), 5–7. https://doi.org/10.1002/cad.20453

Dashti, H. S., Jones, S. E., Wood, A. R., Lane, J. M., van Hees, V. T., Wang, H., Rhodes, J. A., Song, Y., Patel, K., Anderson, S. G., Beaumont, R. N., Bechtold, D. A., Bowden, J., Cade, B. E., Garaulet, M., Kyle, S. D., Little, M. A., Loudon, A. S., Luik, A. I., … Saxena, R. (2019). Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. Nature Communications, 10(1), 1–12. https://doi.org/10.1038/s41467-019-08917-4

De Wolff, M. S., & Van IJzendoorn, M. H. (1997). Sensitivity and Attachment: A Meta-Analysis on Parental Antecedents of Infant Attachment. Child Development, 68(4), 571–591.

Demange, P. A., Malanchini, M., Mallard, T. T., Biroli, P., Cox, S. R., Grotzinger, A. D., Tucker-Drob, E. M., Abdellaoui, A., Arseneault, L., van Bergen, E., Boomsma, D. I., Caspi, A., Corcoran, D. L., Domingue, B. W., Harris, K. M., Ip, H. F., Mitchell, C., Moffitt, T. E., Poulton, R., … Nivard, M. G. (2021). Investigating the genetic architecture of noncognitive skills using GWAS-by-subtraction. Nature Genetics, 53(1), 35–44. https://doi.org/10.1038/s41588-020-00754-2

Deneault, A.-A., Bakermans-Kranenburg, M. J., Groh, A. M., Fearon, P. R. M., & Madigan, S. (2021). Child-father attachment in early childhood and behavior problems: A meta-analysis. Child & Adolescent Development, 2021, 43–66.

Deneault, A.-A., Hammond, S. I., & Madigan, S. (2023). A meta-analysis of child–parent attachment in early childhood and prosociality. Developmental Psychology, 59(2), 236–255. https://doi.org/10.1037/dev0001484

Dobewall, H., Savelieva, K., Seppälä, I., Knafo-Noam, A., Hakulinen, C., Elovainio, M., Keltikangas-Järvinen, L., Pulkki-Råback, L., Raitakari, O. T., Lehtimäki, T., & Hintsanen, M. (2019). Gene-environment correlations in parental emotional warmth and intolerance: Genome-wide analysis over two generations of the Young Finns Study. Journal of Child Psychology and Psychiatry, 60(3), 277–285. https://doi.org/10.1111/jcpp.12995

Doherty, A., Smith-Byrne, K., Ferreira, T., Holmes, M. V., Holmes, C., Pulit, S. L., & Lindgren, C. M. (2018). GWAS identifies 14 loci for device-measured physical activity and sleep duration. Nature Communications, 9(1), 5257. https://doi.org/10.1038/s41467-018-07743-4

Dressendörfer, R. A., Kirschbaum, C., Rohde, W., Stahl, F., & Strasburger, C. J. (1992). Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. The Journal of Steroid Biochemistry and Molecular Biology, 43(7), 683–692. https://doi.org/10.1016/0960-0760(92)90294-S

Dudbridge, F. (2013). Power and Predictive Accuracy of Polygenic Risk Scores. PLoS Genetics, 9(3), e1003348. https://doi.org/10.1371/journal.pgen.1003348

Duncan, L. E., & Keller, M. C. (2011). A Critical Review of the First 10 Years of Candidate Gene-by-Environment Interaction Research in Psychiatry. American Journal of Psychiatry, 168(10), 1041–1049. https://doi.org/10.1176/appi.ajp.2011.11020191

Dykas, M. J., Woodhouse, S. S., Cassidy, J., & Waters, H. S. (2006). Narrative assessment of attachment representations: Links between secure base scripts and adolescent attachment. Attachment & Human Development, 8(3), 221–240. https://doi.org/10.1080/14616730600856099

E

Egeland, B. (1990). 24 months tool coding manual. Project STEEP-revised 1990 from mother-child project scales.

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & Ijzendoorn, M. H. V. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. Development and Psychopathology, 23, 7–28.

El-Sheikh, M., Buckhalt, J. A., Keller, P. S., & Granger, D. A. (2008). Children's Objective and Subjective Sleep Disruptions: Links With Afternoon Cortisol Levels. Health Psychology, 27(1), 26–33. https://doi.org/10.1037/0278-6133.27.1.26

Euser, S., Bakermans-Kranenburg, M. J., Van den Bulk, B. G., Linting, M., Damsteegt, R. C., Vrijhof, C. I., Van Wijk, I. C., Crone, E. A., & Van IJzendoorn, M. H. (2016). Efficacy of the Video-Feedback Intervention to Promote Positive Parenting and Sensitive Discipline in Twin Families (VIPP-Twins): Study Protocol for a Randomized Controlled Trial. BMC Psychology, 4(33). https://doi.org/10.1186/s40359-016-0139-y

Euser, S., Bosdriesz, J. R., Vrijhof, C. I., van den Bulk, B. G., van Hees, D., de Vet, S. M., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2020). How Heritable are Parental Sensitivity and Limit-Setting? A Longitudinal Child-Based Twin Study on Observed Parenting. Child Development, 91(6), 2255–2269. https://doi.org/10.1111/cdev.13365

Euser, S., Vrijhof, C. I., Van den Bulk, B. G., Vermeulen, R., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2021). Video-feedback promotes sensitive limit-setting in parents of twin preschoolers: A randomized controlled trial. BMC Psychology, 9(1), 46. https://doi.org/10.1186/s40359-021-00548-z

Eythorsdottir, D. Y., Frederiksen, P., Larsen, S. C., Olsen, N. J., & Heitmann, B. L. (2020). Associations between objective measures of physical activity, sleep and stress levels among preschool children. BMC Pediatrics, 20(1), 258. https://doi.org/10.1186/s12887-020-02108-7

F

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. Behavior Research Methods, 39, 175–191

Fernandez-Mendoza, J., Vgontzas, A. N., Calhoun, S. L., Vgontzas, A., Tsaoussoglou, M., Gaines, J., Liao, D., Chrousos, G. P., & Bixler, E. O. (2014). Insomnia symptoms, objective sleep duration and hypothalamic-pituitary-adrenal activity in children. European Journal of Clinical Investigation, 44(5), 493–500. https://doi.org/10.1111/eci.12263

Fonagy, P. (2022). Maternal Representations of Attachment during Pregnancy Predict the Organization of Infant-Mother Attachment at One Year of Age.

Fraley, C. R. (2002). Attachment Stability From Infancy to Adulthood: Meta-Analysis and Dynamic Modeling of Developmental Mechanisms. Personality and Social Psychology Review, 6(2), 123–151. https://doi.org/10.1207/S15327957PSPR0602_03

G

Gao, W., Stalder, T., Foley, P., Rauh, M., Deng, H., & Kirschbaum, C. (2013). Quantitative analysis of steroid hormones in human hair using a column-switching LC–APCI–MS/MS assay. Journal of Chromatography B, 928, 1–8. https://doi.org/10.1016/j.jchromb.2013.03.008

Gehrman, P. R., Ghorai, A., Goodman, M., McCluskey, R., Barilla, H., Almasy, L., Roenneberg, T., & Bucan, M. (2019). Twin-based heritability of actimetry traits. Genes, Brain and Behavior, e12569. https://doi.org/10.1111/gbb.12569

Genderson, M. R., Rana, B. K., Panizzon, M. S., Grant, M. D., Toomey, R., Jacobson, K. C., Xian, H., Cronin-Golomb, A., Franz, C. E., Kremen, W. S., & Lyons, M. J. (2013). Genetic and environmental influences on sleep quality in middle-aged men: A twin study. Journal of Sleep Research, 22(5), 519–526. https://doi.org/10.1111/jsr.12048

Ghassabian, A., Steenweg-de Graaff, J., Peeters, R. P., Ross, H. A., Jaddoe, V. W. V., Hofman, A., Verhulst, F. C., White, T., & Tiemeier, H. (2014). Maternal urinary iodine concentration in pregnancy and children's cognition: Results from a population-based birth cohort in an iodine-sufficient area. BMJ Open, 47(e005520). https://doi.org/doi: 10.1136/bmjopen-2014-005520

Gillath, O., Hart, J., Noftle, E. E., & Stockdale, G. D. (2009). Development and validation of a state adult attachment measure (SAAM). Journal of Research in Personality, 43(3), 362–373. https://doi.org/10.1016/j.jrp.2008.12.009

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry, 38(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Granger, D. A., Hibel, L. C., Fortunato, C. K., & Kapelewski, C. H. (2009). Medication effects on salivary cortisol: Tactics and strategy to minimize impact in behavioral and developmental science. Psychoneuroendocrinology, 34(10), 1437–1448. https://doi.org/10.1016/j.psyneuen.2009.06.017

Green, B. L., Furrer, C., & McAllister, C. (2007). How Do Relationships Support Parenting? Effects of Attachment Style and Social Support on Parenting Behavior in an At-Risk Population. American Journal of Community Psychology, 40(1–2), 96–108. https://doi.org/10.1007/s10464-007-9127-y

Groh, A. M., Fearon, R. M. P., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., & Roisman, G. I. (2017). Attachment in the Early Life Course: Meta-Analytic Evidence for Its Role in Socioemotional Development. Child Development Perspectives, 11(1), 70–76.

Groh, A. M., Roisman, G. I., Booth-LaForce, C., Fraley, R. C., Owen, M. T., Cox, M. J., & Burchinal, M. R. (2014). Stability of Attachment Security from Infancy to late Adolescence. Monographs of the Society for Research in Child Development, 79(3), 51–66. https://doi.org/10.1111/mono.12113

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M., Deary, I. J., Koellinger, P. D., Harden, K. P., Nivard, M. G., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nature Human Behaviour, 3(5), 513–525. https://doi.org/10.1038/s41562-019-0566-x

Gunnar, M. R. (2017). Social Buffering of Stress in Development: A Career Perspective. Perspectives on Psychological Science, 12(3), 355–373. https://doi.org/10.1177/1745691616680612

Gustafsson, P.A., Gustafsson, P.E., Anckarsatar, H., Lichtenstein, P., Ljung, T., Nelson, N., Larsson, H., 2011. Heritability of cortisol regulation in children. Twin Res. Hum. Genet. 14, 553–561.

Gutman, L., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. A literature review. Education Endowment Foundation. https://doi.org/10.1007/s10654-016-0224-9 [pii]

H

Hamaker, E. L., Mulder, J. D., & van IJzendoorn, M. H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development. Developmental Cognitive Neuroscience, 46, 100867. https://doi.org/10.1016/j.dcn.2020.100867

Harden, K. P., & Koellinger, P. D. (2020). Using genetics for social science. Nature Human Behaviour, 4(6), 567–576. https://doi.org/10.1038/s41562-020-0862-5

Hatzinger, M., Brand, S., Perren, S., Von Wyl, A., Stadelmann, S., von Klitzing, K., & Holsboer-Trachsler, E. (2013). In pre-school children, sleep objectively assessed via sleep-EEGs remains stable over 12 months and is related to psychological functioning, but not to cortisol secretion. Journal of Psychiatric Research, 47(11), 1809–1814. https://doi.org/10.1016/j.jpsychires.2013.08.007

Hawkins, E., Madigan, S., Moran, G., & Pederson, D. R. (2015). Mediating and moderating processes underlying the association between maternal cognition and infant attachment. Journal of Applied Developmental Psychology, 39, 24–33. https://doi.org/10.1016/j.appdev.2015.04.001

Heckman, J. J., & Masterov, D. V. (2007). The Productivity Argument for Investing in Young Children. Review of Agricultural Economics, 29(3), 446–493.

Hesse, E. (2016). Measurement of Individual Differences in Adult Attachment. In Handbook of Attachment: Theory, research, and clinical applications (Third). Guilford Press.

Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., Di Angelantonio, E., Roberts, D. J., Xueyi, S., Davies, G., Liewald, D. C. M., Porteous, D. J., Hayward, C., Butterworth, A. S., McIntosh, A. M., Gale, C. R., & Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. Nature Communications, 10(1), 5741. https://doi.org/10.1038/s41467-019-13585-5

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., Hazen, N., Herman, J., Katz, E. S., Kheirandish-Gozal, L., Neubauer, D. N., O'Donnell, A. E., Ohayon, M., Peever, J., Rawding, R., Sachdeva, R. C., Setters, B., Vitiello, M. V., Ware, J. C., & Adams Hillard, P. J. (2015). National Sleep Foundation's sleep time duration recommendations: Methodology and results summary. Sleep Health, 1(1), 40–43. https://doi.org/10.1016/j.sleh.2014.12.010

Hostinar, C. E., Sullivan, R. M., & Gunnar, M. R. (2014). Psychobiological mechanisms underlying the social buffering of the hypothalamic–pituitary–adrenocortical axis: A review of animal models and human studies across development. Psychological Bulletin, 140(1), 256–282. https://doi.org/10.1037/a0032671

Hu, L., & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification. Psychological Methods, 3(4), 424.

Hublin, C., Partinen, M., Koskenvuo, M., & Kaprio, J. (2013). Genetic factors in evolution of sleep length—A longitudinal twin study in Finnish adults. Journal of Sleep Research, 22(5), 513–518. https://doi.org/10.1111/jsr.12051

I

Inderkum, A. P., & Tarokh, L. (2018). High heritability of adolescent sleep–wake behavior on free, but not school days: A long-term twin study. Sleep, 41(3). https://doi.org/10.1093/sleep/zsy004

J

Jackson, E., Galvin, J., Warrier, V., Baron-Cohen, S., Luo, S., Dunbar, R. I., Proctor, H., Lee, E., & Richards, G. (2022). Evidence of assortative mating for theory of mind via facial expressions but not language. Journal of Social and Personal Relationships, 39(12), 3660–3679. https://doi.org/10.1177/02654075221106451

Jaddoe, V. W. V., Bakker, R., van Duijn, C. M., van der Heijden, A. J., Lindemans, J., Mackenbach, J. P., Moll, H. A., Steegers, E. A. P., Tiemeier, H., Uitterlinden, A. G., Verhulst, F. C., & Hofman, A. (2007). The Generation R Study Biobank: A resource for epidemiological studies in children and their parents. European Journal of Epidemiology, 22(12), 917–923. https://doi.org/10.1007/s10654-007-9209-z

Juffer, F., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2017). Pairing attachment theory and social learning theory in video-feedback intervention to promote positive parenting. Current Opinion in Psychology, 15, 189–194. https://doi.org/10.1016/j.copsyc.2017.03.012

K

Kagan, J. (2007). A Trio of Concerns. Perspectives on Psychological Science, 2(4), 361–376. https://doi.org/10.1111/j.1745-6916.2007.00049.x

Kagan, J., Reznick, J. S., & Snidman, N. (1987). The Physiology and Psychology of Behavioral Inhibition in Children. Child Developmen, 58(6), 1459–1473. https://doi.org/10.2307/1130685

Kahan, B. C., Jairath, V., Doré, C. J., & Morris, T. P. (2014). The risks and rewards of covariate adjustment in randomized trials: An assessment of 12 outcomes from 8 studies. Trials, 15(1), 139. https://doi.org/10.1186/1745-6215-15-139

Keers, R., Coleman, J. R. I., Lester, K. J., Roberts, S., Breen, G., Thastum, M., Bögels, S., Schneider, S., Heiervang, E., Meiser-Stedman, R., Nauta, M., Creswell, C., Thirlwall, K., Rapee, R. M., Hudson, J. L., Lewis, C., Plomin, R., & 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

Kenny, D. A. (2017). MedPower: An interactive tool for the estimation of power in tests of mediation.

Kenny, D. A., & McCoach, D. B. (2003). Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. Structural Equation Modeling: A Multidisciplinary Journal, 10(3), 333–351. https://doi.org/10.1207/S15328007SEM1003_1

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary Cortisol in Psychoneuroendocrine Research: Recent Developments and Applications. Psychoneuroendocrinology, 19, 313–333. https://doi.org/10.1016/0306-4530(94)90013-2

Klahr, A. M., & Burt, S. A. (2014). Elucidating the etiology of individual differences in parenting: A meta-analysis of behavioral genetic research. Psychological Bulletin, 140(2), 544–586. https://doi.org/10.1037/a0034205

Knafo, A., & Jaffee, S. R. (2013). Gene–environment correlation in developmental psychopathology. Development and Psychopathology, 25(1), 1–6. https://doi.org/10.1017/S0954579412000855

Kok, R., Linting, M., Bakermans-Kranenburg, M., Van IJzendoorn, M. H., Jaddoe, V. W. V., Hofman, A., Verhulst, F. C., & Tiemeier, H. (2013). Maternal Sensitivity and Internalizing Problems: Evidence from Two Longitudinal Studies in Early Childhood. Child Psychiat Hum D, 44(6), 751–765. https://doi.org/doi:10.1007/s10578-013-0369-7

Kooijman, M. N., Kruithof, C. J., van Duijn, C. M., Duijts, L., Franco, O. H., van IJzendoorn, M. H., de Jongste, J. C., Klaver, C. C. W., van der Lugt, A., Mackenbach, J. P., Moll, H. A., Peeters, R. P., Raat, H., Rings, E. H. H. M., Rivadeneira, F., van der Schroeff, M. P., Steegers, E. A. P., Tiemeier, H., Uitterlinden, A. G., … Jaddoe, V. W. V. (2016). The Generation R Study: Design and cohort update 2017. European Journal of Epidemiology, 31(12), 1243–1264. https://doi.org/10.1007/s10654-016-0224-9

Koopman-Verhoeff, M. E., Serdarevic, F., Kocevska, D., Bodrij, F. F., Mileva-Seitz, V. R., Reiss, I., Hillegers, M. H. J., Tiemeier, H., Cecil, C. A. M., Verhulst, F. C., & Luijk, M. P. C. M. (2019). Preschool family irregularity and the development of sleep problems in childhood: A longitudinal study. Journal of Child Psychology and Psychiatry, jcpp.13060. https://doi.org/10.1111/jcpp.13060

Krapohl, E., Patel, H., Newhouse, S., Curtis, C. J., von Stumm, S., Dale, P. S., Zabaneh, D., Breen, G., O'Reilly, P. F., & Plomin, R. (2018). Multi-polygenic score approach to trait prediction. Molecular Psychiatry, 23(5), 1368–1374. https://doi.org/10.1038/mp.2017.163

Kretschmer, T. (2023). Parenting is genetically influenced: What does that mean for research into child and adolescent social development? Social Development, 32(1), 3–16. https://doi.org/10.1111/sode.12633

Krishnakumar, A., & Buehler, C. (2000). Interparental Conflict and Parenting Behaviors: A Meta-Analytic Review. Family Relations, 49(1), 25–44. https://doi.org/10.1111/j.1741-3729.2000.00025.x

L

Lamballais, S., Jansen, P. R., Labrecque, J. A., Ikram, M. A., & White, T. (2021). Genetic scores for adult subcortical volumes associate with subcortical volumes during infancy and childhood. Human Brain Mapping, 42(6), 1583–1593. https://doi.org/10.1002/hbm.25292

Lane, J. M., Liang, J., Vlasac, I., Anderson, S. G., Bechtold, D. A., Bowden, J., Emsley, R., Gill, S., Little, M. A., Luik, A. I., Loudon, A., Scheer, F. A. J. L., Purcell, S. M., Kyle, S. D., Lawlor, D. A., Zhu, X., Redline, S., Ray, D. W., Rutter, M. K., & Saxena, R. (2017). Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. Nature Genetics, 49(2), 274–281. https://doi.org/10.1038/ng.3749

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., … Turley, P. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genetics, 50(8), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Lee, S. S., Chronis-Tuscano, A., Keenan, K., Pelham, W. E., Loney, J., Van Hulle, C. A., Cook, E. H., & Lahey, B. B. (2010). Association of maternal dopamine transporter genotype with negative parenting: Evidence for gene x environment interaction with child disruptive behavior. Molecular Psychiatry, 15(5), 548–558. https://doi.org/10.1038/mp.2008.102

Leerkes, E. M., Blankson, A. N., & O'Brien, M. (2009). Differential Effects of Maternal Sensitivity to Infant Distress and Nondistress on Social-Emotional Functioning: Sensitivity to Infant Distress. Child Development, 80(3), 762–775. https://doi.org/10.1111/j.1467-8624.2009.01296.x

Lemery-Chalfant, K., Clifford, S., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2018). Genetic moderation of the effects of the Family Check-Up intervention on children's internalizing symptoms: A longitudinal study with a racially/ethnically diverse sample. Development and Psychopathology, 30(5), 1729–1747. https://doi.org/10.1017/S095457941800127X

Lemola, S., Perkinson-Gloor, N., Hagmann-von Arx, P., Brand, S., Holsboer-Trachsler, E., Grob, A., & Weber, P. (2015). Morning cortisol secretion in school-age children is related to the sleep pattern of the preceding night. Psychoneuroendocrinology, 52(1), 297–301. https://doi.org/10.1016/j.psyneuen.2014.12.007

Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S., & Hirschhorn, J. N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nature Genetics, 33(2), 177–182. https://doi.org/10.1038/ng1071

Lopez-Minguez, J., Morosoli, J. J., Madrid, J. A., Garaulet, M., & Ordoñana, J. R. (2017). Heritability of siesta and night-time sleep as continuously assessed by a circadian-related integrated measure. Scientific Reports, 7(1), 12340. https://doi.org/10.1038/s41598-017-12460-x

Lovejoy, M. C., Graczyk, P. A., O'Hare, E., & Neuman, G. (2000). Maternal Depression and Parenting Behavior: A Meta-analytic review. Clinical Psychology Review, 20(5), 561–592.

M

Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. Journal of Clinical Epidemiology, 110, 63–73. https://doi.org/10.1016/j.jclinepi.2019.02.016

Main, M., & Hesse, E. (1990). Parents' unresolved traumatic experiences are related to infant disorganized attachment status: Is frightened/frightening parental behavior the linking mechanism? In Attachment in the preschool years: Theory, research, and intervention (p. 161182). University of Chicago Press.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in Infancy, Childhood, and Adulthood: A Move to the Level of Representation. Monographs of the Society for Research in Child Development, 50(1/2), 66. https://doi.org/10.2307/3333827

Marceau, K., Abel, E. A., Duncan, R. J., Moore, P. J., Leve, L. D., Reiss, D., Shaw, D. S., Natsuaki, M., Neiderhiser, J. M., & Ganiban, J. M. (2019). Longitudinal Associations of Sleep Duration, Morning and Evening Cortisol, and BMI During Childhood. Obesity, 27(4), 645–652. https://doi.org/10.1002/oby.22420

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. Structural Equation Modeling: A Multidisciplinary Journal, 11(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Martins, R. C., Blumenberg, C., Tovo-Rodrigues, L., Gonzalez, A., & Murray, J. (2020). Effects of parenting interventions on child and caregiver cortisol levels: Systematic review and meta-analysis. BMC Psychiatry, 20(1), 370. https://doi.org/10.1186/s12888-020-02777-9

Medina-Gomez, C., Felix, J. F., Estrada, K., Peters, M. J., Herrera, L., Kruithof, C. J., Duijts, L., Hofman, A., van Duijn, C. M., Uitterlinden, A. G., Jaddoe, V. W. V., & Rivadeneira, F. (2015). Challenges in conducting genome-wide association studies in highly admixed multi-ethnic populations: The Generation R Study. European Journal of Epidemiology, 30(4), 317–330. https://doi.org/10.1007/s10654-015-9998-4

Meehl, P. E. (1990). Appraising and Amending Theories: The Strategy of Lakatosian Defense and Two Principles that Warrant It. Psychological Inquiry, 1(2), 108–141. https://doi.org/10.1207/s15327965pli0102_1

Merckelbach, H., Smeets, T., & Jelicic, M. (2009). Experimental simulation: Type of malingering scenario makes a difference. Journal of Forensic Psychiatry & Psychology, 20(3), 378–386. https://doi.org/10.1080/14789940802456686

Mileva-Seitz, V. R., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2016). Genetic mechanisms of parenting. Hormones and Behavior, 77, 211–223. https://doi.org/10.1016/j.yhbeh.2015.06.003

Mills-Koonce, W., Gariepy, J.-L., Sutton, K., & Cox, M. (2008). Changes in maternal sensitivity across the first three years: Are mothers from different attachment dyads differentially influenced by depressive symptomatology? Attachment & Human Development, 10(3), 299–317. https://doi.org/doi:10.1080/14616730802113612

Min, J. L., Hemani, G., Hannon, E., Dekkers, K. F., Castillo-Fernandez, J., Luijk, R., Carnero-Montoro, E., Lawson, D. J., Burrows, K., Suderman, M., Bretherick, A. D., Richardson, T. G., Klughammer, J., Iotchkova, V., Sharp, G., Khleifat, A. A., Shatunov, A., Iacoangeli, A., McArdle, W. L., … Relton, C. L. (2021). Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nature Genetics, 53(September 2021), 1311–1321. https://doi.org/10.1038/s41588-021-00923-x

Miner, J. L., & Clarke-Stewart, K. A. (2008). Trajectories of externalizing behavior from age 2 to age 9: Relations with gender, temperament, ethnicity, parenting, and rater. Developmental Psychology, 44(3), 771–786. https://doi.org/10.1037/0012-1649.44.3.771

Moffitt, T. E., Caspi, A., & Rutter, M. (2006). Measured Gene-Environment Interactions in Psychopathology: Concepts, Research Strategies, and Implications for Research, Intervention, and Public Understanding of Genetics. Perspectives on Psychological Science, 1(1), 5–27. https://doi.org/10.1111/j.1745-6916.2006.00002.x

Moore, R. Y. (2007). Suprachiasmatic nucleus in sleep–wake regulation. Sleep Medicine, 8, 27–33. https://doi.org/10.1016/j.sleep.2007.10.003

Müller, M. B., & Wurst, W. (2004). Getting closer to affective disorders: The role of CRH receptor systems. Trends in Molecular Medicine, 10(8), 409–415. https://doi.org/10.1016/j.molmed.2004.06.007

Musca, S. C., Kamiejski, R., Nugier, A., Méot, A., Er-Rafiy, A., & Brauer, M. (2011). Data with Hierarchical Structure: Impact of Intraclass Correlation and Sample Size on Type-I Error. Frontiers in Psychology, 2. https://doi.org/10.3389/fpsyg.2011.00074

Muthén, L. K., & Muthén, B. O. (2017). Mplus User's Guide (8th ed.). Muthén & Muthén.

N

Nagel, M., Speed, D., Sluis, S., & Østergaard, S. D. (2020). Genome-wide association study of the sensitivity to environmental stress and adversity neuroticism cluster. Acta Psychiatrica Scandinavica, 141(5), 476–478. https://doi.org/10.1111/acps.13155

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended Structural Equation and Statistical Modeling. Psychometrika, 81(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8

Neuhauser, A. (2018). Predictors of maternal sensitivity in at-risk families. Early Child Dev Care, 188(2), 126–142. https://doi.org/doi:10.1080/03004430.2016.1207065

Neumann, A., Jolicoeur-Martineau, A., Szekely, E., Sallis, H. M., O'Donnel, K., Greenwood, C. M. T., Levitan, R., Meaney, M. J., Wazana, A., Evans, J., & Tiemeier, H. (2022). Combined polygenic risk scores of different psychiatric traits predict general and specific psychopathology in childhood. Journal of Child Psychology and Psychiatry, 63(6), 636–645. https://doi.org/10.1111/jcpp.13501

Newman, D. A. (2014). Missing Data: Five Practical Guidelines. Organizational Research Methods, 17(4), 372–411. https://doi.org/10.1177/1094428114548590

Nivison, M. D., Facompré, C. R., Raby, K. L., Simpson, J. A., Roisman, G. I., & Waters, T. E. A. (2021). Childhood abuse and neglect are prospectively associated with scripted attachment representations in young adulthood. Development and Psychopathology, 33(4), 1143–1155. https://doi.org/10.1017/S0954579420000528

O

Oakley, R. H., & Cidlowski, J. A. (2013). The biology of the glucocorticoid receptor: New signaling mechanisms in health and disease. Journal of Allergy and Clinical Immunology, 132(5), 1033–1044. https://doi.org/10.1016/j.jaci.2013.09.007

O'Farrelly, C., Barker, B., Watt, H., Babalis, D., Bakermans-Kranenburg, M., Byford, S., Ganguli, P., Grimås, E., Iles, J., Mattock, H., McGinley, J., Phillips, C., Ryan, R., Scott, S., Smith, J., Stein, A., Stevens, E., van IJzendoorn, M., Warwick, J., & Ramchandani, P. (2021). A video-feedback parenting intervention to prevent enduring behaviour problems in at-risk children aged 12–36 months: The Healthy Start, Happy Start RCT. Health Technology Assessment, 25(29), 1–84. https://doi.org/10.3310/hta25290

Okbay, A., Wu, Y., & Wang, N. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. Nature Genetics, 54(4). https://doi.org/doi:10.1038/s41588-022-01016-z

Ouellet-Morin, I., Brendgen, M., Girard, A., Lupien, S. J., Dionne, G., Vitaro, F., & Boivin, M. (2016). Evidence of a unique and common genetic etiology between the CAR and the remaining part of the diurnal cycle: A study of 14 year-old twins. Psychoneuroendocrinology, 66, 91–100. https://doi.org/10.1016/j.psyneuen.2015.12.022

P

Parlak, O., Keene, S. T., Marais, A., Curto, V. F., & Salleo, A. (2018). Molecularly selective nanoporous membrane-based wearable organic electrochemical device for noninvasive cortisol sensing. Science Advances, 4(7), 1–10. https://doi.org/10.1126/sciadv.aar2904

Pesonen, A. K., Kajantie, E., Heinonen, K., Pyhälä, R., Lahti, J., Jones, A., Matthews, K. A., Eriksson, J. G., Strandberg, T., & Räikkönen, K. (2012). Sex-specific associations between sleep problems and hypothalamic-pituitary-adrenocortical axis activity in children. Psychoneuroendocrinology, 37(2), 238–248. https://doi.org/10.1016/j.psyneuen.2011.06.008

Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. Developmental Psychology, 53(5), 873–932. https://doi.org/10.1037/dev0000295

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-Environment Interaction and Correlation in the Analysis of Human Behavior. Psychological Bulletin, 84(2), 309–322.

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. Nature Reviews Genetics, 19(3), 148–159. https://doi.org/10.1038/nrg.2017.104

Poehlmann-Tynan, J., Engbretson, A., Vigna, A. B., Weymouth, L. A., Burnson, C., Zahn-Waxler, C., Kapoor, A., Gerstein, E. D., Fanning, K. A., & Raison, C. L. (2020). Cognitively-Based Compassion Training for parents reduces cortisol in infants and young children. Infant Mental Health Journal, 41(1), 126–144. https://doi.org/10.1002/imhj.21831

Powell, K. (2021). The broken promise that undermines human genome research. Nature, 590(7845), 198–201. https://doi.org/10.1038/d41586-021-00331-5

Privé, F., Arbel, J., & Vilhjálmsson, B. J. (2021). LDpred2: Better, faster, stronger. Bioinformatics, 36(22–23), 5424–5431. https://doi.org/10.1093/bioinformatics/btaa1029

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics, 81(3), 559–575. https://doi.org/10.1086/519795

R

Räikkönen, K., Matthews, K. A., Pesonen, A.-K., Pyhälä, R., Paavonen, E. J., Feldt, K., Jones, A., Phillips, D. I. W., Seckl, J. R., Heinonen, K., Lahti, J., Komsi, N., Järvenpää, A.-L., Eriksson, J. G., Strandberg, T. E., & Kajantie, E. (2010). Poor Sleep and Altered Hypothalamic-Pituitary-Adrenocortical and Sympatho-Adrenal-Medullary System Activity in Children. The Journal of Clinical Endocrinology & Metabolism, 95(5), 2254–2261. https://doi.org/10.1210/jc.2009-0943

R Core Team. (2017). R: A language and environment for statistical computing.

Reul, J. M., & Holsboer, F. (2002). Corticotropin-releasing factor receptors 1 and 2 in anxiety and depression. Current Opinion in Pharmacology, 2(1), 23–33. https://doi.org/10.1016/S1471-4892(01)00117-5

Rippe, R. C. A., Noppe, G., Windhorst, D. A., Tiemeier, H., van Rossum, E. F. C., Jaddoe, V. W. V., Verhulst, F. C., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & van den Akker, E. L. T. (2016). Splitting hair for cortisol? Associations of socio-economic status, ethnicity, hair color, gender and other child characteristics with hair cortisol

Rietveld, M.J.H., Van der Valk, J.C., Bongers, I.L., Stroet, T.M., E, S.P., Boomsma, D.I., 2000. Zygosity diagnosis in young twins by parental report. Twin Res. 3, 134–141.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How Much Does Education Improve Intelligence? A Meta-Analysis. Psychological Science, 29(8), 1358–1369. https://doi.org/doi.org/10.1177/0956797618774253

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

Rohner, R. P., Khaleque, A., & Cournoyer, D. E. (2008). Parental Acceptance-Rejection: Theory, Methods, Cross-Cultural Evidence, and Implications. Ethos, 33(3), 299–334.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2). https://doi.org/10.18637/jss.v048.i02

Rothbart, M. K. (1981). Measurement of Temperament in Infancy. Child Development, 52, 569–578.

Runze, J., Pappa, I., Van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2022). Conduct Problems and Hair Cortisol Concentrations Decrease in School-Aged Children after VIPP-SD: A Randomized Controlled Trial in Two Twin Cohorts. International Journal of Environmental Research and Public Health, 19(22), 15026. https://doi.org/10.3390/ijerph192215026

Runze, J., Van IJzendoorn, M. H., Vrijhof, C. I., & Bakermans-Kranenburg, M. J. (2022). Replicating a randomized trial with video-feedback to promote positive parenting in parents of school-aged twins. Journal of Family Psychology, 36(4), 490–501. https://doi.org/10.1037/fam0000961

Russell, E., Koren, G., Rieder, M., & Van Uum, S. (2012). Hair cortisol as a biological marker of chronic stress: Current status, future directions and unanswered questions. Psychoneuroendocrinology, 37(5), 589–601. https://doi.org/10.1016/j.psyneuen.2011.09.009

S

Sadeh, A. (2015). Sleep Assessment Methods. Monogr Soc Res Child, 80, 33–49. https://doi.org/10.1093/acprof:oso/9780195395754.003.0015

Sadeh, A., Sharkey, M., & Carskadon, M. A. (1994). Activity-Based Sleep-Wake Identification: An Empirical Test of Methodological Issues. Sleep, 17(3), 201–207. https://doi.org/10.1093/sleep/17.3.201

Sadeh, A., Tikotzky, L., & Kahn, M. (2014). Sleep in infancy and childhood: Implications for emotional and behavioral difficulties in adolescence and beyond. Current Opinion in Psychiatry, 27(6), 453–459. https://doi.org/10.1097/YCO.0000000000000109

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing Structural Equation Models or Detection of Misspecifications? Structural Equation Modeling: A Multidisciplinary Journal, 16(4), 561–582. https://doi.org/10.1080/10705510903203433

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., … Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nature Genetics, 50(7), 912–919. https://doi.org/10.1038/s41588-018-0152-6

Shaw, D. S., Galán, C. A., Lemery-Chalfant, K., Dishion, T. J., Elam, K. K., Wilson, M. N., & Gardner, F. (2019). Trajectories and Predictors of Children's Early-Starting Conduct Problems: Child, Family, Genetic, and Intervention Effects. Development and Psychopathology, 31(5), 1911–1921. https://doi.org/10.1017/S0954579419000828

Shewark, E. A., Ramos, A. M., Liu, C., Ganiban, J. M., Fosco, G., Shaw, D. S., Reiss, D., Natsuaki, M. N., Leve, L. D., & Neiderhiser, J. M. (2021). The role of child negative emotionality in parenting and child adjustment: Gene–environment interplay. Journal of Child Psychology and Psychiatry, 62(12), 1453–1461. https://doi.org/10.1111/jcpp.13420

Simmons, J. G., Azpitarte, F., Roost, F. D., Dommers, E., Allen, N. B., Havighurst, S., & Haslam, N. (2019). Correlates of hair cortisol concentrations in disadvantaged young children. Stress and Health, 35(1), 104–111. https://doi.org/10.1002/smi.2842

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Psychological Science, 22(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Slagt, M., Dubas, J. S., Deković, M., & van Aken, M. A. G. (2016). Differences in sensitivity to parenting depending on child temperament: A meta-analysis. Psychological Bulletin, 142(10), 1068–1110. https://doi.org/10.1037/bul0000061

Sletten, T. L., Rajaratnam, S. M. W., Wright, M. J., Zhu, G., Naismith, S., Martin, N. G., & Hickie, I. (2013). Genetic and Environmental Contributions to Sleep-Wake Behavior in 12-Year-Old Twins. Sleep, 36(11), 1715–1722. https://doi.org/10.5665/sleep.3136

Stalder, T., & Kirschbaum, C. (2012). Analysis of cortisol in hair – State of the art and future directions. Brain, Behavior, and Immunity, 26(7), 1019–1029. https://doi.org/10.1016/j.bbi.2012.02.002

Stalder, T., Steudte-Schmiedgen, S., Alexander, N., Klucken, T., Vater, A., Wichmann, S., Kirschbaum, C., & Miller, R. (2017). Stress-related and basic determinants of hair cortisol in humans: A meta-analysis. Psychoneuroendocrinology, 77, 261–274. https://doi.org/10.1016/j.psyneuen.2016.12.017

Steele, R. D., Waters, T. E. A., Bost, K. K., Vaughn, B. E., Truitt, W., Waters, H. S., Booth-LaForce, C., & Roisman, G. I. (2014). Caregiving antecedents of secure base script knowledge: A comparative analysis of young adult attachment representations. Developmental Psychology, 50(11), 2526–2538. https://doi.org/10.1037/a0037992

Steptoe, A., van Jaarsveld, C. H. M., Semmler, C., Plomin, R., & Wardle, J. (2009). Heritability of daytime cortisol levels and cortisol reactivity in children. Psychoneuroendocrinology, 34(2), 273–280. https://doi.org/10.1016/j.psyneuen.2008.09.006

Strahler, J., Skoluda, N., Kappert, M. B., & Nater, U. M. (2017). Simultaneous Measurement of Salivary Cortisol and Alpha-Amylase: Application and Recommendations. Neuroscience & Biobehavioral Reviews, 83, 657–677. https://doi.org/10.1016/j.neubiorev.2017.08.015

T

Taraban, L., & Shaw, D. S. (2018). Parenting in context: Revisiting Belsky's classic process of parenting model in early childhood. Developmental Review, 48, 55–81. https://doi.org/10.1016/j.dr.2018.03.006

Taylor, M. J., Gregory, A. M., Freeman, D., & Ronald, A. (2015). Do sleep disturbances and psychotic-like experiences in adolescence share genetic and environmental influences? Journal of Abnormal Psychology, 124(3), 674–684. https://doi.org/10.1037/abn0000057

Tiemeier, H., Velders, F. P., Szekely, E., Roza, S. J., Dieleman, G., Jaddoe, V. W. V., Uitterlinden, A. G., White, T. J. H., Bakermans-Kranenburg, M. J., Hofman, A., Van IJzendoorn, M. H., Hudziak, J. J., & Verhulst, F. C. (2012). The Generation R Study: A Review of Design, Findings to Date, and a Study of the 5-HTTLPR by Environmental Interaction From Fetal Life Onward. Journal of the American Academy of Child & Adolescent Psychiatry, 51(11), 1119-1135.e7. https://doi.org/10.1016/j.jaac.2012.08.021

Tierney, N., Cook, D., McBain, M., & Fay, C. (2022). naniar: Data Structures, Summaries, and Visualisations for Missing Data (R package version 0.6.1.9000). https://github.com/njtierney/naniar

Torvik, F. A., Eilertsen, E. M., Hannigan, L. J., Cheesman, R., Howe, L. J., Magnus, P., Reichborn-Kjennerud, T., Andreassen, O. A., Njølstad, P. R., Havdahl, A., & Ystrom, E. (2022). Modeling assortative mating and genetic similarities between partners, siblings, and in-laws. Nature Communications, 13(1), 1108. https://doi.org/10.1038/s41467-022-28774-y

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38(1), 1–10. https://doi.org/10.1007/BF02291170

U

Unternaehrer, E., Cost, K. T., Bouvette-Turcot, A. A., Gaudreau, H., Massicotte, R., Dhir, S. K., Hari Dass, S. A., O'Donnell, K. J., Gordon-Green, C., Atkinson, L., Levitan, R. D., Wazana, A., Steiner, M., Lydon, J. E., Clark, R., Fleming, A. S., Meaney, M. J., & MAVAN Research Team. (2019). Dissecting maternal care: Patterns of maternal parenting in a prospective cohort study. J Neuroendocrinol, 31(9). https://doi.org/doi:ARTN e12784 10.1111/jne.12784

V

Valcan, D. S., Davis, H., & Pino-Pasternak, D. (2018). Parental Behaviours Predicting Early Childhood Executive Functions: A Meta-Analysis. Educational Psychology Review, 30(3), 607–649. https://doi.org/10.1007/s10648-017-9411-9

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Van Dam, M., & van IJzendoorn, M. H. (1988). Measuring Attachment Security: Concurrent and Predictive Validity of the Parental Attachment Q-set. The Journal of Genetic Psychology, 149(4), 447–457.

Van Der Mark, I. L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2002). The role of parenting, attachment, and temperamental fearfulness in the prediction of compliance in toddler girls. British Journal of Developmental Psychology, 20(3), 361–378. https://doi.org/10.1348/026151002320620299

Van der Meulen, M., Steinbeis, N., Achterberg, M., Van IJzendoorn, M. H., & Crone, E. A. (2018). Heritability of Neural Reactions to Social Exclusion and Prosocial Compensation in Middle Childhood. Developmental Cognitive Neuroscience, 34, 42–52. https://doi.org/10.1016/j.dcn.2018.05.010

Van Doesum, K., Hosman, C., Riksen-Walraven, J., & Hoefnagels, C. (2007). Correlates of depressed mothers' sensitivity toward their infants: The role of maternal, child, and contextual characteristics. J Am Acad Child Psy, 46(6), 747–756. https://doi.org/doi:DOI 10.1097/CHI.0b013e318040b272

Van IJzendoorn, M. H. (1995). Adult Attachment Representations, Parental Responsiveness, and Infant Attachment: A Meta-Analysis on the Predictive Validity of the Adult Attachment Interview. Psychological Bulletin, 117(3), 387–403. https://doi.org/0033-2909/95/$3.00

Van IJzendoorn, M. H. (1984). Answers without Questions: A Note on Response Style in Questionnaires. Perceptual and Motor Skills, 59(3), 827–831. https://doi.org/10.2466/pms.1984.59.3.827

Van IJzendoorn, M. H., Schuengel, C., Wang, Q., & Bakermans-Kranenburg, M. J. (2023). Improving parenting, child attachment, and externalizing behaviors: Meta-analysis of the first 25 randomized controlled trials on the effects of Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline. Development and Psychopathology, 1–16. https://doi.org/10.1017/S0954579421001462

Van Keulen, B. J., Dolan, C. V., Andrew, R., Walker, B. R., Hulshoff Pol, H. E., Boomsma, D. I., Rotteveel, J., & Finken, M. J. J. (2020). Heritability of Cortisol Production and Metabolism Throughout Adolescence. The Journal of Clinical Endocrinology & Metabolism, 105(2), 443–452. https://doi.org/10.1210/clinem/dgz016

Van Lenten, S. A., & Doane, L. D. (2016). Examining multiple sleep behaviors and diurnal salivary cortisol and alpha-amylase: Within- and between-person associations. Psychoneuroendocrinology, 68, 100–110. https://doi.org/10.1016/j.psyneuen.2016.02.017

Vaughn, B. E., Veríssimo, M., Coppola, G., Bost, K. K., Shin, N., McBride, B., Krzysik, L., & Korth, B. (2006). Maternal attachment script representations: Longitudinal stability and associations with stylistic features of maternal narratives. Attachment & Human Development, 8(3), 199–208. https://doi.org/10.1080/14616730600856024

Vázquez-Palacios, G., Retana-Márquez, S., Bonilla-Jaime, H., & Velázquez-Moctezuma, J. (2001). Further definition of the effect of corticosterone on the sleep–wake pattern in the male rat. Pharmacology Biochemistry and Behavior, 70(2–3), 305–310. https://doi.org/10.1016/S0091-3057(01)00620-7

Verhage, M. L., Fearon, R. M. P., Schuengel, C., van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Madigan, S., Roisman, G. I., Oosterman, M., Behrens, K. Y., Wong, M. S., Mangelsdorf, S., Priddis, L. E., Brisch, K.-H., & The Collaboration on Attachment Transmission Synthesis. (2018). Examining Ecological Constraints on the Intergenerational Transmission of Attachment Via Individual Participant Data Meta-analysis. Child Development, 89(6), 2023–2037. https://doi.org/10.1111/cdev.13085

Verhage, M. L., Schuengel, C., Madigan, S., Fearon, R. M. P., Oosterman, M., Cassibba, R., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2016). Narrowing the transmission gap: A synthesis of three decades of research on intergenerational transmission of attachment. Psychological Bulletin, 142(4), 337–366. https://doi.org/10.1037/bul0000038

Visscher, P. M. (2004). Power of the classical twin design revisited. Twin Research, 7, 505–512.

Vrijhof, C. I., van den Bulk, B. G., Overgaauw, S., Lelieveld, G., Engels, R. C. M. E., & van IJzendoorn, M. H. (2016). The Prosocial Cyberball Game: Compensating for social exclusion and its associations with empathic concern and bullying in adolescents. Journal of Adolescence, 52(1), 27–36. https://doi.org/10.1016/j.adolescence.2016.07.005

W

Walker, S. A., Double, K. S., Kunst, H., Zhang, M., & MacCann, C. (2022). Emotional intelligence and attachment in adulthood: A meta-analysis. Personality and Individual Differences, 184, 111174. https://doi.org/10.1016/j.paid.2021.111174

Ward, J., Lyall, L., Cullen, B., Strawbridge, R. J., Zhu, X., Stanciu, I., Aman, A., Niedzwiedz, C. L., Anderson, J., Bailey, M. E. S., Lyall, D. M., & Pell, J. (2022). A Genome-Wide Association Analysis of Happiness: Consistent Genetic Effects Across the Lifespan and Across Genetic Ancestries in Multiple Cohorts [Preprint]. Genetics. https://doi.org/10.1101/2022.04.05.487098

Warrier, V., Kwong, A. S. F., Luo, M., Dalvie, S., Croft, J., Sallis, H. M., Baldwin, J., Munafò, M. R., Nievergelt, C. M., Grant, A. J., Burgess, S., Moore, T. M., Barzilay, R., McIntosh, A., van IJzendoorn, M. H., & Cecil, C. A. M. (2021). Gene–environment correlations and causal effects of childhood maltreatment on physical and mental health: A genetically informed approach. The Lancet Psychiatry, 8(5), 373–386. https://doi.org/10.1016/S2215-0366(20)30569-1

Waters, E., & Deane, K. E. (1985). Defining and Assessing Individual Differences in Attachment Relationships: Q-Methodology and the Organization of Behavior in Infancy and Early Childhood. Monographs of the Society for Research in Child Development, 50(1/2), 41. https://doi.org/10.2307/3333826

Waters, H. S., & Waters, E. (2006). The attachment working models concept: Among other things, we build script-like representations of secure base experiences. Attachment & Human Development, 8(3), 185–197. https://doi.org/10.1080/14616730600856016

Waters, T. E. A., Bosmans, G., Vandevivere, E., Dujardin, A., & Waters, H. S. (2015). Secure base representations in middle childhood across two Western cultures: Associations with parental attachment representations and maternal reports of behavior problems. Developmental Psychology, 51(8), 1013–1025. https://doi.org/10.1037/a0039375

Waters, T. E. A., Facompré, C. R., Dagan, O., Martin, J., Johnson, W. F., Young, E. S., Shankman, J., Lee, Y., Simpson, J. A., & Roisman, G. I. (2021). Convergent validity and stability of secure base script knowledge from young adulthood to midlife. Attachment & Human Development, 23(5), 740–760. https://doi.org/10.1080/14616734.2020.1832548

Waters, T. E. A., Facompré, C. R., Dujardin, A., Van De Walle, M., Verhees, M., Bodner, N., Boldt, L. J., & Bosmans, G. (2019). Taxometric Analysis of Secure Base Script Knowledge in Middle Childhood Reveals Categorical Latent Structure. Child Development, 90(3), 694–707. https://doi.org/10.1111/cdev.13229

Waters, T. E. A., Facompré, C. R., Van de Walle, M., Dujardin, A., De Winter, S., Heylen, J., Santens, T., Verhees, M., Finet, C., & Bosmans, G. (2019). Stability and change in secure base script knowledge during middle childhood and early adolescence: A 3-year longitudinal study. Developmental Psychology, 55(11), 2379–2388. https://doi.org/10.1037/dev0000798

Waters, T. E., & Roisman, G. I. (2019). The secure base script concept: An overview. Current Opinion in Psychology, 25, 162–166. https://doi.org/10.1016/j.copsyc.2018.08.002

Waters, T. E. A., Ruiz, S. K., & Roisman, G. I. (2017). Origins of Secure Base Script Knowledge and the Developmental Construction of Attachment Representations. Child Development, 88(1), 198–209. https://doi.org/10.1111/cdev.12571

Waters, T. E. A., Yang, R., Finet, C., Verhees, M. W. F. T., & Bosmans, G. (2022). An empirical test of prototype and revisionist models of attachment stability and change from middle childhood to adolescence: A 6-year longitudinal study. Child Development, 93(1), 225–236. https://doi.org/10.1111/cdev.13672

Wechsler, D. (1991). Wechsler Intelligence Scale for Children—Third Edition (WISC-III). The Psychological Corporation.

Wertz, J., Belsky, J., Moffitt, T. E., Belsky, D. W., Harrington, H. L., Avinun, R., Poulton, R., Ramrakha, S., & Caspi, A. (2019). Genetics of nurture: A test of the hypothesis that parents' genetics predict their observed caregiving. Developmental Psychology, 55(7), 1461–1472. https://doi.org/10.1037/dev0000709

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd, L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA From Mothers and Children to Study Parental Investment in Children's Educational Attainment. Child Development, 91(5), 1745–1761. https://doi.org/10.1111/cdev.13329

Windhorst, D. A., Mileva-Seitz, V. R., Linting, M., Hofman, A., Jaddoe, V. W. V., Verhulst, F. C., Tiemeier, H., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2015). Differential susceptibility in a developmental perspective: DRD4 and maternal sensitivity predicting externalizing behavior: G × E Effects of DRD4 and Sensitivity Over Time. Developmental Psychobiology, 57(1), 35–49. https://doi.org/10.1002/dev.21257

Wildman, R. W., & Wildman, R. W. (1999). The Detection of Malingering. Psychological Reports, 84, 386–388.

Witte, A. M., Runze, J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2023). Parents' secure base script knowledge predicts observed sensitive caregiving and discipline toward twin children. Journal of Family Psychology. https://doi.org/10.1037/fam0001091

Wulff, J. N., & Ejlskov, L. (2017). Multiple Imputation by Chained Equations in Praxis: Guidelines and Review. Electronic Journal on Business Research Methods, 15(1), 41–56.

X

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. Behavior Research Methods, 51(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

Y

Yanagida, T. (2023). _misty: Miscellaneous Functions. [R package version 0.4.10]. https://CRAN.R-project.org/package=misty

Yang, J., Lee, S., Goddard, M., & Visscher, P. (2011). GCTA: A Tool for Genome-wide Complex Trait Analysis. Am J Hum Genet, 88(1), 76–82. https://doi.org/doi:10.1016/j.ajhg.2010.11.011

Yang, J., & Qiu, M. (2016). The impact of education on income inequality and intergenerational mobility. China Economic Review, 37, 110–125. https://doi.org/10.1016/j.chieco.2015.12.009

Z

Zagorsky, J. L. (2007). Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. Intelligence, 35(5), 489–501. https://doi.org/10.1016/j.intell.2007.02.003

Zhang, X., Widaman, K., & Belsky, J. (2021). Beyond orchids and dandelions: Susceptibility to environmental influences is not bimodal. Development and Psychopathology, 1–13. https://doi.org/10.1017/S0954579421000821

Zeiders, K. H., Doane, L. D., & Adam, E. K. (2011). Reciprocal Relations Between Objectively Measured Sleep Patterns and Diurnal Cortisol Rhythms in Late Adolescence. Journal of Adolescent Health, 48(6), 566–571. https://doi.org/10.1016/j.jadohealth.2010.08.012

Author contribution statements

Chapter 2

Jana Runze: data curation, methodology, formal analysis, investigation, visualization, writing - original draft, writing - review and editing
Marinus H. van IJzendoorn: funding acquisition, methodology, project administration, supervision, validation, writing - original draft, writing - review and editing

Chapter 3

Jana Runze: conceptualization, methodology, formal analysis, writing - original draft, writing - review and editing, visualization,
Marinus H. van IJzendoorn: conceptualization, methodology, writing - review and editing, supervision, funding acquisition
Charlotte Cecil: investigation, resources, writing - review and editing
Marian J. Bakermans-Kranenburg: conceptualization, methodology, writing - review and editing, supervision, funding acquisition
Irene Pappa: conceptualization, methodology, formal analysis, data curation, writing - original draft, writing - review and editing

Chapter 4

Jana Runze: conceptualization, methodology, formal analysis, investigation, data curation, writing - original draft, writing - review and editing, visualization,
Saskia Euser: methodology, investigation, data curation
Mirjam Oosterman: writing - review and editing, supervision
Conor V. Dolan: methodology, validation, formal analysis, writing - original draft, writing - review and editing
M. Elisabeth Koopman-Verhoeff: methodology, writing - review and editing
Marian J. Bakermans-Kranenburg: methodology, writing - review and editing, supervision, funding acquisition

Chapter 5

Jana Runze: conceptualization, methodology, formal analysis, investigation, data curation, writing - original draft, writing - review and editing, visualization,
Marinus H. van IJzendoorn: conceptualization, methodology, writing - review and editing, supervision, funding acquisition
Claudia I. Vrijhof: investigation, writing - original draft, writing - review and editing
Marian J. Bakermans-Kranenburg: conceptualization, methodology, writing - review and editing, supervision, funding acquisition

Chapter 6

Jana Runze: conceptualization, methodology, formal analysis, investigation, data curation, writing - original draft, writing - review and editing, visualization
Irene Pappa: formal analysis, data curation, writing - review and editing
Marinus H. van IJzendoorn: conceptualization, methodology, validation, writing - review and editing, funding acquisition
*Marian J. Bakermans-Kranenbur*g: conceptualization, methodology, investigation, writing - review and editing, supervision, funding acquisition

Chapter 7

Jana Runze: conceptualization, methodology, formal analysis, investigation, writing - original draft
Annemieke M. Witte: writing - original draft, writing - review and editing
Marinus H. van IJzendoorn: methodology, writing - review and editing, funding acquisition, resources, supervision
Mirjam Oosterman: writing - review and editing, supervision
Marian J. Bakermans-Kranenburg: methodology, writing - review and editing, funding acquisition, resources, supervision

Summary

Parents, and their multifaceted relationship with and behavior towards their children play a major role in a child's development. The development of children does not only hold significance for the children themselves, but also for society since children grow up to be active members of our society. Problems in the psychobiological development of children have been linked to (mental) health difficulties later in life and subsequent personal, social and economic costs. Preventing the onset of (mental) health difficulties by supporting adaptive child development is therefore highly desired.

Within the broad context of parenting, the main questions of this thesis were what the (genetic) determinants of parental reports and behavior are and in how far parents and genes are determinants of children's attachment and psychobiology. These questions were answered using two studies: the Leiden Consortium on Individual Development (L-CID) and the Generation R study (GenR). L-CID is a longitudinal cohort-sequential study including two cohorts of parents and their twin children: an early childhood cohort with 238 families and a middle childhood cohort with 257 families. GenR is a population-based prospective cohort study with an enrolment of 9778 pregnant mothers.

In **Chapter 2**, I examined the potential role of genes in response bias of parents and their twin children. We examined the genetic predisposition for the acquiescence response bias in 257 families with twins. Both parents and twins completed a modified Wildman Symptoms Questionnaire and provided salivary DNA samples. From published Genome-Wide Association Studies (GWAS) we derived polygenic score (PGS) algorithms for indicators of socio-economic status (educational attainment and income) and for IQ, which were applied to the GWAS results of our participants' DNA. Higher polygenic scores for educational attainment and income but not for IQ predicted less acquiescence response bias in parents and children.

In **Chapter 3,** I investigated whether genetic predispositions to socio-economic and cognitive factors (i.e. educational attainment, IQ, and income)

can predict sensitive parenting in mothers of the GenR study. A higher maternal PGS of educational attainment (PGS-EA) predicted higher observed sensitivity in toddlerhood and early childhood. Child PGS-EA was significantly associated with maternal sensitivity in early childhood, and the effect of maternal PGS-EA was no longer significant when correcting for child PGS-EA. A latent factor of PGSs based on educational attainment, intelligence (IQ), and income showed similar results. These polygenic scores might be associated with maternal cognitive and behavioral skills that help shape parenting. Maternal PGSs predicted observed sensitivity over and above the maternal phenotypes, showing an additional role for PGSs and yielding results that are more easily interpreted in a causal manner.

In **Chapter 4**, I investigated the heritability of sleep and cortisol, in order to examine the potential for environmental (i.e., parenting) effects on sleep and cortisol of their children. The heritability of sleep duration, sleep efficiency and wake episodes were 52%, 45%, and 55%, respectively. Common environmental factors played no significant role. High genetic correlations between sleep duration and sleep efficiency and high genetic correlations between sleep efficiency and wake episodes were found. Shared environmental (29%), and unique environmental factors (53%) explained the variance in morning cortisol levels. Because the sleep and cortisol measures were found to be uncorrelated, we did not consider genetic and environmental contributions to the association between the sleep and cortisol measures. Our findings indicate that sleep duration, sleep efficiency, and wake episodes in children are mostly impacted by genetic factors and by unique environmental factors.

In **Chapter 5**, I examined whether a randomized controlled trial in the form of a brief attachment-based video-feedback intervention (VIPP-SD) increased parental sensitivity and sensitive discipline in parents of school-aged twin children. In our sample with older children, the VIPP-SD did not significantly change observed parental sensitivity or sensitive discipline in the intervention group compared to the control group. The VIPP-SD did improve parents' attitude toward sensitivity, but not toward discipline. Intervention effects were not moderated by temperamental reactivity of the parents, providing no support for the differential susceptibility

hypothesis.

Chapter 6 expanded the research about VIPP-SD by investigating its effect on children's hair cortisol levels and conduct problems. In the current pre-registered trial, we randomly assigned 40% of n = 445 families with school-aged twin children to the intervention group. The VIPP-SD was successful in decreasing both children's conduct problems and hair cortisol levels. Effects were not moderated by available child PGS of differential susceptibility. We conclude that a brief, home-based video-feedback parenting intervention can decrease child behavior problems and affect the child's stress-related neuroendocrine system as assessed with hair cortisol.

In **Chapter 7**, I examined whether the transmission of attachment representation is mediated by parental sensitivity and sensitive discipline and whether children are differentially susceptible to parental sensitivity and sensitive discipline. Multilevel structural equation models showed that parental sensitivity, but not parental secure base script knowledge or parental sensitive discipline predicted children's attachment. This association was moderated by child temperamental discomfort: lower levels of parental sensitivity predicted less secure child attachment in children with higher discomfort, but not in children with lower discomfort.

In **Chapter 8**, findings of the previous chapters were discussed. Future research can make significant steps in parenting research by incorporating genome-wide association studies of observed parenting measures and making use of polygenic scores in the quest for domain-specific differential susceptibility.

Publications

Runze, J., Bakermans-Kranenburg, M.J., Cecil, C., Van IJzendoorn, M.H., & Pappa, I. (2023). The Polygenic and Reactive Nature of Observed Parenting. *Genes, Brain, and Behavior, 22*. doi: 10.1111/gbb.12874

Runze, J. & Van IJzendoorn, M.H. (2023). Response Bias is Genetically Biased: Another argument for Kagan's Philippic against Questionnaires in Developmental Psychology. *Developmental Psychology*. doi: 10.1037/dev0001614. Epub ahead of print. PMID: 37902678.

Witte, A. M., **Runze,** J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2023). Parents' secure base script knowledge predicts observed sensitive caregiving and discipline toward twin children. Journal of Family Psychology. https://doi.org/10.1037/fam0001091

Runze, J., Pappa, E., Van IJzendoorn M.H., & Bakermans-Kranenburg, M.J. (2022). Conduct problems and hair cortisol concentrations decrease in school-aged children after VIPP-SD: A randomized controlled trial in two twin cohorts. International Journal of Environmental Research and Public Health, 19(22), 15026. https://doi.org/10.3390/ijerph192215026

Runze, J., Marten, F., & te Brinke, L., (2022). The effect of exposure to social-media coverage of the Russo-Ukrainian war on stress symptoms in Dutch adolescents (Pre-print)

Runze, J., Van IJzendoorn, M. H., Vrijhof, C. I., & Bakermans-Kranenburg, M. J. (2022). Replicating a randomized trial with video-feedback to promote positive parenting in parents of school-aged twins. Journal of Family Psychology. 2022 Jan 27. doi: 10.1037/fam0000961

Toenders, Y.J., Van der Cruijsen, R., **Runze**, J., Van de Groep, S., Wieringa, L. & Crone, E.A. (2022). Mood fluctuations during development and their relation to brain development and sleep (Pre-print)

Runze, J., Euser, S., Oosterman, M., Dolan, C. V., Koopman-Verhoeff, M. E., & Bakermans-Kranenburg, M. J. (2021). Actigraphic sleep and cortisol in middle childhood: A multivariate behavioral genetics model. Comprehensive Psychoneuroendocrinology, 8, 100094. https://doi.org/10.1016/j.cpnec.2021.100094

Becker, S., Hopps, D., Owens, G., **Runze**, J., Morris, S., Bainbridge, V., & Wylie, M. (2017). Co-creating by degrees exploring experiences of co design. Digital or Visual Products

Under review/In revision

Witte, A.M., **Runze,** J., Bakermans-Kranenburg, M.J., & Van IJzendoorn, M.H. (2024). Effectiveness of VIPP-SD: Secure Base Script Knowledge as an Outcome and Moderator [In revision]

Runze, J., Witte, A.M., Van IJzendoorn, M.H., Oosterman, M., & Bakermans-Kranenburg, M.J. (2024). Differential Susceptibility in the Intergenerational Transmission of Secure Base Script Knowledge? [Under review]

Runze, J., Overbeek, G., Luik, A.I., & ten Have, M. (2024). Does Child Abuse Predict a Population Segment with Large Economic Burden [Submitted]

Curriculum Vitae

Born on August 10, 1992 in Gummersbach, Germany, Jana Runze completed her high school education at Gymnasium Moltkestrasse in Gummersbach in 2011. She then pursued an apprenticeship as business specialist in marketing and communication at MediaCom in Düsseldorf, from which she graduated with distinction in January 2014. After spending six months in British Columbia, Canada, she studied Psychology at the University of Twente in Enschede. During her bachelor's degree, Jana spent a semester studying at Teesside University in Middlesbrough in the UK. She graduated cum laude and as valedictorian in 2017. After receiving her Bachelor of Sciences, Jana enrolled in the two-year Research Master in Behavioral Sciences at Radboud University in Nijmegen. She conducted her internship at the Developmental Psychobiology Lab under the supervision of prof.dr. Carolina de Werth. In her thesis, she investigated the association between child attachment and children's telomere length as well as the potential moderating role of behavioral control. In 2019, Jana obtained her research masters (cum laude) and began her PhD at the Vrije Universiteit Amsterdam, as part of the Leiden Consortium on Individual Development. She was supervised by prof. dr. Marian Bakermans-Kranenburg and dr. Mirjam Oosterman and investigated the genetic and environmental, as well as their interactional effects on parent-child relationship and child development. In 2022, Jana was awarded the FGB Talent Fund which enabled her to visit the ALSPAC lab in Bristol. Since May 2023, Jana works as a postdoctoral researcher at the Research Institute of Child Development and Education of the University of Amsterdam under the supervision of prof.dr. Geertjan Overbeek.

PhD Portfolio

Name PhD student:	Jana Runze
Department:	Clinical Child & Family Studies
University:	Vrije Universiteit Amsterdam
PhD period:	August 2019 – July 2023
Promotor:	Prof. dr. Marian Bakermans-Kranenburg
Copromotor:	Dr. Mirjam Oosterman

1. PhD Training	**Year**
Courses	
Attachment: Strange situation procedure (D)	2023
Career orientation	2023
Attachment: Strange situation procedure (ABC)	2023
Gene finding: Genome-Wide Association Studies and beyond	2022
University Teaching qualification (UTQ/BKO)	2022
Modeling techniques for longitudinal analyses	2022
Multilevel SEM	2022
Advanced Multilevel Analysis	2021
Behavior Genetics	2020
Writing a Data Management Plan	2019
Research Integrity	2019
Open Science	2019
Coding Sensitive Discipline	2019
(Inter)national conferences	
ECDP, Turku, Finland	2023
ISDP, Utrecht, The Netherlands	2023
SRCD, Salt Lake City, USA	2023
ISDP, San Diego, USA	2022
IAC, Lisbon, Portugal	2022
BAPS, Leuven, Belgium	2022
VNOP, Utrecht, The Netherlands	2022
LEARN! Conference, Amsterdam, The Netherlands	2022
L-CID Conference, Leiden, The Netherlands	2022
ISDP, Chicago, USA	2021
SRCD, online	2021
ISPNE, online	2020

Meetings & workshops	
Amsterdam Public Health (APH) Annual Meeting, Amsterdam	2022
VU Show & Share, Amsterdam,	2022
VU Data Conversations, Amsterdam,	2022
CID meeting, Utrecht	2022
VNOP Research days, online	2022
VNOP Research days, online	2021
CID autumn meeting, Utrecht	2019
Dutch Neuroscience Meeting, Lunteren	2019
2. Teaching activities	
Supervision of 7 Master theses	2019-2023
Supervision of 1 Bachelor thesis	2019-2023
Supervision of 18 research internships	2019-2023
Teaching observational measure for caregiver-child interaction	2019-2021
Coordinating Master course: “Master thesis studio”	2021-2022
Co-coordinating Master course “Intervention in Research & Practice”	2022
Guest lecturing in Master Pedagogical Sciences	2019-2023
3. Other activities	
Member of section head application committee	2022
Conference co-organizer for the L-CID congress	2022
PhD Co-representative	2022
Organizer of bi-weekly research meetings	2021-2023
Member of the PhD Education Committee	2020-2023
Member of the Forum for Young Scientists	2020-2023
Reviewer for several journals	2021-
4. Grants & prices	
EADP: Travel fellowship (€250)	2023
ISDP: Postdoc Member Travel award ($500)	2023
VU Amsterdam, FGB: Open Science Award (€200)	2023
VU Vereniging subsidy for Forum for Young Scientists (€7000)	2023
Open Science certificate of Open Science Community Amsterdam	2023
Society for Research on Child Development: Travel grant ($300)	2023
ISDP: Student Abstract Award ($325)	2022
VU Amsterdam, FGB: Talent Fund (€4500)	2022
VU Amsterdam, FGB: Open Science Award (€200)	2022
SEAS: Bursary for the International Attachment Conference (€500)	2022
ISDP: Student Member Abstract Award (registration costs)	2021
APH Junior Award (conference costs)	2020

Acknowledgements

First, I would like to thank all parents and their twin children of the L-CID study. Without your participation, there would have been no data to investigate within my PhD project. I am also thankful to the participants of the GenR study and the Urban Rotterdam project as well as the research groups who shared their data.

My promotors, **Marian** and **Mirjam**, thank you both for your support and trust in me. **Marian**, I am glad to have gotten the chance of working with you. At each step of my PhD, you encouraged and supported me, whatever my plans were. You helped me forming a solid foundation for my future career. I couldn't have asked for a better supervisor. **Mirjam**, you have been my link to the department. You supported me and made me feel comfortable during my PhD.

Many other researchers have been part of my journey. **Nic Timpson** & your whole team at the University of Bristol, thank you for welcoming me in your team! **Rien**, thank you for challenging my views, being critical and sharing new methods and insights. I enjoy working with you and am grateful for all the opportunities you provided. **Irene**, I am happy that you have joined the L-CID team. You have enriched my knowledge and challenged my views with your data-driven, clinical viewpoints. **Annemieke**, I am so glad you joined L-CID. I just love how similar we are in so many things, I couldn't imagine sitting in an AirBnb during a conference week with someone else talking about missing our dogs and trash TV.

I also want to thank the whole **L-CID team** for creating this warm atmosphere and being such nice, helpful colleagues. **Eveline**, thank you for caring even though I am not one of your ducklings. **Judith & Lisanne**, you were my partners in crime (home visits!) and I wouldn't have lasted a week without you when I started my PhD.

Also thanks to my new research home, the **preventive youth care** group at the UvA. I am grateful that everyone of the team welcomed me so warmly! Thanks specifically to **Geertjan**, for offering me this great postdoc position and being

enthusiastic about my research ideas.

Friends are the family we choose. **Maike**, obwohl wir uns vielleicht einmal im Jahr sehen und du mittlerweile in Südtirol wohnst, ist es doch jedes Mal so als hätten wir uns gestern das letzte Mal getroffen. Ich hoffe, dass das immer so bleiben wird und freue mich darauf, dich schnell besuchen zu kommen. **Nathan** and **Olivia**. I am so glad to have met you when I moved to Diemen (-Zuid!). Together, we survived the first year of the pandemic and I enjoy every time we come together for coffee, sports, borrel or a nice dinner, even though only one of us still lives in Diemen-Zuid.

Finja, ich bin froh, dass wir Freunde sind. Wir haben so viel gemeinsam, unsere Affinität für Research, unsere Liebe zu Hunden, gutes Essen und schlechte Serien, unsere Gym Erfahrungen und unsere introvertierte Extraversion. Während des kompletten Masters (und unwissentlich auch im Bachelor) warst du an meiner Seite und hast mir geholfen, mich auf den PhD vorzubereiten. Ich weiß nicht ob ich ohne dich an dem Punkt wäre an dem ich gerade stehe.

Leonie, danke, dass du meine Freundin bist! Du bist so loyal zu jedem deiner Freunde und ich schätze es sehr dass du immer Zeit für mich findest in deinem vollen Termibplaner. Du zeigst mir eine andere Perspektive fürs Leben und ich bin froh dafür. Du hast mich bei jedem Schritt seit des Masters unterstützt, und wortwörtlich bei jedem Schritt meines ersten Marathons. Das werde ich nie vergessen und ich hoffe, dass noch viele Marathons, Städtetrips, Kaffee & Kuchen Pausen in süßen Cafés und Drinks in Irish Pubs (oder auf dem Anna Fest) folgen werden.

Nicole, wir kennen uns schon so viele Jahre. Wer hätte damals, mit 14 (oder 13? oder bei der Taufe?), gedacht, dass ich dir in einer Doktorarbeit eine Danksagung schreibe? Ich bin sehr dankbar für deine Loyalität, deine Akzeptanz, deinen Humor und dass du immer für mich da warst in den letzten 17 oder mehr Jahren. Ich hoffe, dass unsere Freundschaft noch ganz viele Jahre weiterhin so bestehen bleibt (Schauen wir mal, was wird was wird) und ich liebs dass du jetzt jedem der es hören will (oder auch nicht) stolz erzählen kannst, dass ich Dr. Jana Runze bin. Und noch

ein paar Worte an deine Mama: **Margot**, du bist wie eine zweite Mutter für mich! Es freut mich so sehr, dass auch du stolz auf mich bist und ich danke dir für deine warme, liebevolle Art.

Familie Verhoeff & Breevaart, jullie kreeg ik er gratis bij Jan bij en daar ben ik uitermate blij om. Ik vind het ontzettend leuk hoe geïnteresseerd jullie in mijn onderzoek zijn.

Ines, ich glaube es gibt für mich nichts Besseres, als eine Schwester, die prahlt mit meinen Erfolgen, weil sie unglaublich stolz ist (trotz Allem).

Ten slotte, **Jan*.** Je bent mijn grootste steun gewezen tijdens mijn PhD. Ik heb zo vaak mijn praatjes aan je gepresenteerd, om hulp gevraagd in R, Illustrator, Linux en InDesign, ik heb zo vaak gezeurd over van alles en jij hebt altijd geluisterd (of gedaan alsof). Eigenlijk had je in elke acknowledgement moeten staan. Je steunt mij, je moedigt me aan en je bent er voor me, altijd. Samen kunnen we de hele wereld veroveren. Dankjewel, ik hou van je (NB: wat een cheesy tekst dit, sorry hoor!).

*en Lea