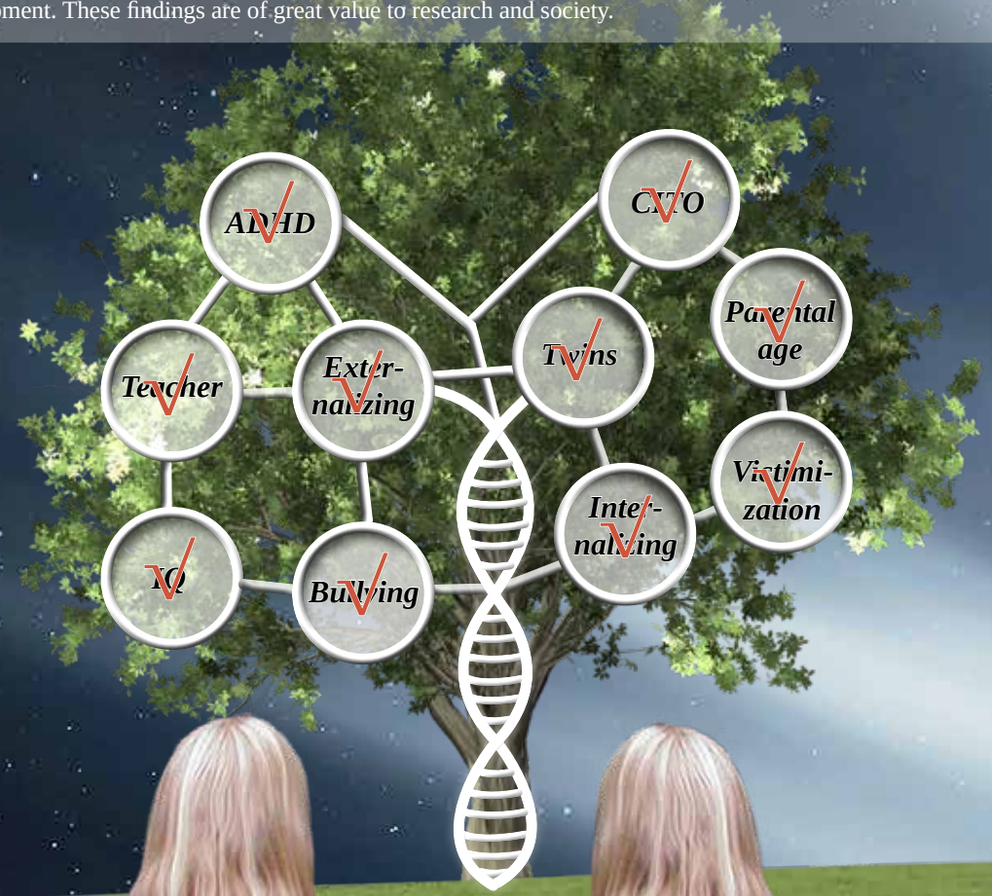
Childhood developmental problems, like being bullied or having attention problems, are common and can have serious lifelong consequences. It is therefore of global concern to investigate what makes some children vulnerable to face difficulties and makes other children resilient. A broad range of candidate factors were investigated by using twin and population cohorts. The studies presented in this thesis focused on two important domains in the field of childhood development: bullying in primary school and the influence of parental age at birth on offspring's development. For the first domain, I investigated the influence of various risk and protective factors on bullying perpetration and bullying victimization. Moreover, I investigated whether bullying runs in families due to shared genetic liabilities or due to the shared family environment. I concluded that some children are at increased risk for being a perpetrator or victim of bullying mostly due to their genetic liability. Regarding the second domain, I examined the influence of parental age on children's social-emotional and cognitive development. I employed advanced statistical approaches to four large Dutch cohorts and found that advanced parental age has no negative effects on children's development. These findings are of great value to research and society.

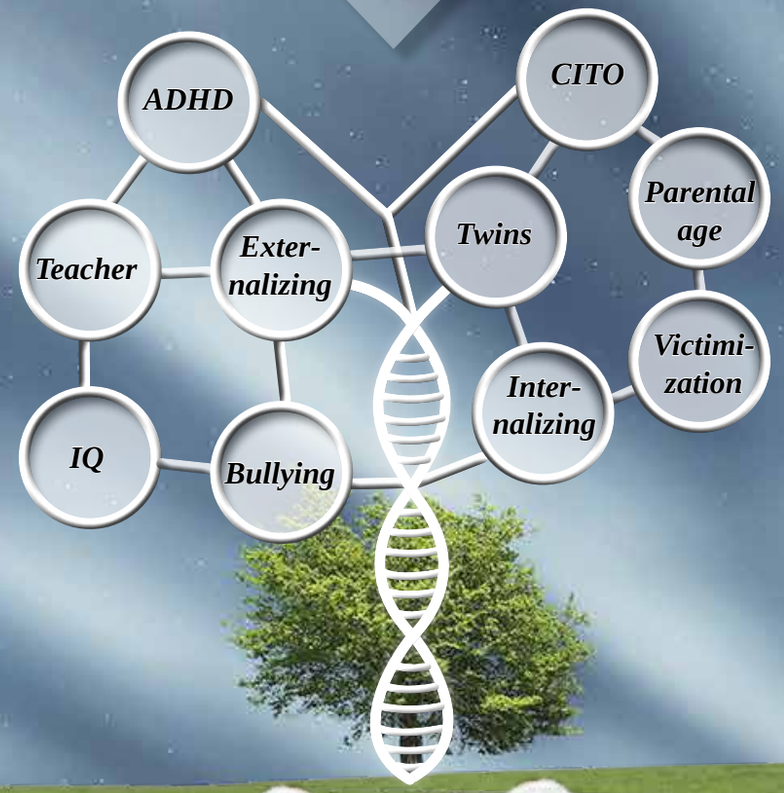


Childhood individual differences

Sabine Veldkamp

Childhood individual differences

Sabine Veldkamp



Risk and protective factors in twin and population cohorts

CHILDHOOD INDIVIDUAL DIFFERENCES

RISK AND PROTECTIVE FACTORS IN TWIN AND POPULATION COHORTS

Sabine Veldkamp

Reading committee
prof.dr. M.E.J. Raijmakers
prof.dr. C. Kemner
prof.dr. M.H.J. Hillegers
prof.dr. S.J.T. Branje
dr. M.H.M. de Moor
dr. T. Kretschmer

Paranymphs
Bas Dobbe
Tom Niemantsverdriet

Acknowledgements
We are grateful to the twin families and the teachers for participation. This research was financially supported by The Consortium on Individual Development (CID). CID is funded through the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant number 024.001.003).

ISBN: 978-90-9032132-5
Printed by: GVO drukkers & vormgevers B.V.
Cover design: A.C. Veldkamp
Layout: A.C. Veldkamp

VRIJE UNIVERSITEIT

CHILDHOOD INDIVIDUAL DIFFERENCES

RISK AND PROTECTIVE FACTORS IN TWIN AND POPULATION COHORTS

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op woensdag 18 september 2019 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Sabine Anne Maria Veldkamp

geboren te Deventer

promotoren:	prof.dr. D.I. Boomsma
	prof.dr. M. Bartels
copromotoren:	dr. E. van Bergen
	dr. L.E.J. de Zeeuw

Voor mijn geliefden

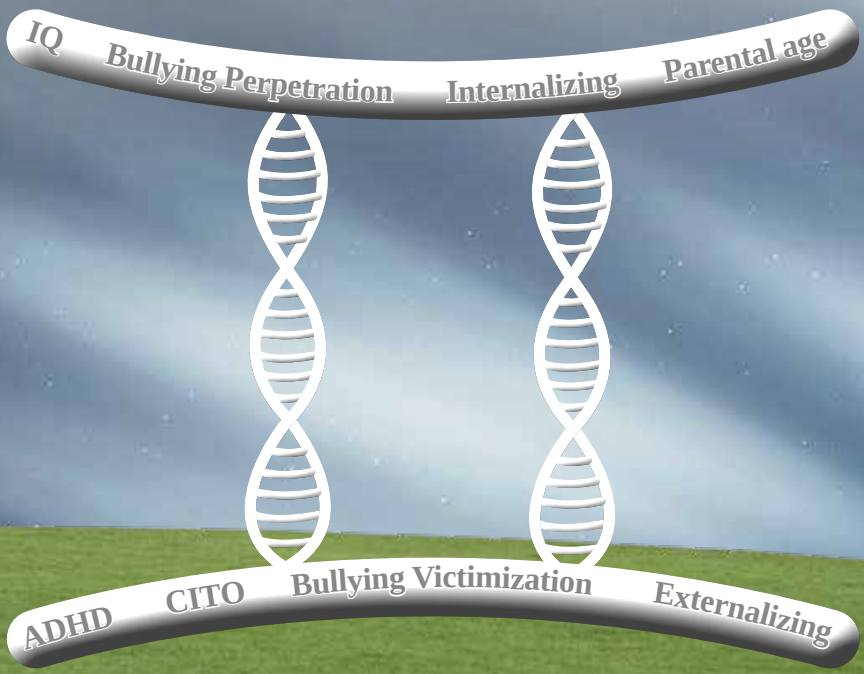
“Optimism is about seeing possibilities without losing reality.”

“Strength grows in the moments when you think you can’t go on, but keep going anyway.”

Contents

Chapter 1

General Introduction



What makes some children vulnerable to face difficulties and what makes others resilient? Differences between children's developmental paths are due to a combination of the children's biological predisposition, the children's characteristics, and the children's rearing environment. In other words, child development is affected by multiple influences: biological, psychological, and environmental factors and the interplay between them. The aim of this thesis is to advance our understanding of why some children thrive and others not. It is of global concern to unravel the causes of individual differences in social-emotional and cognitive development in order to reduce the prevalence of childhood problems and its lifelong impact. In sum, we wondered what makes some children vulnerable and others resilient?

Within this overarching question, this thesis focused on two very important, but poorly understood topics: 1) bullying during primary school, and 2) influences of parental age on child development. For bullying, I looked at influences of risk factors on the prevalence of perpetration and victimization as well as at the etiology of different subtypes of perpetration (the ones who bully), victimization (the victims of bullying), and the association between perpetration and victimization. Regarding parental age influences, I gained unique insight into the extent to which a children's development is influenced by parental age at birth. We focused on the influences of parental age on the children's socio-emotional and cognitive development.

Research on child development so far has been hampered by the traditional boundaries of the research areas involved. This inspired the formation of the Consortium on Individual Development (CID), which brings together researchers from several Dutch universities from a wide range of behavioral and social science disciplines. Bundling expertise may lead to understanding that goes beyond earlier attempts to understand child development. This consortium includes four different work packages (WP) with large existing and new youth cohorts in the Netherlands to study childhood development, see Figure 1.

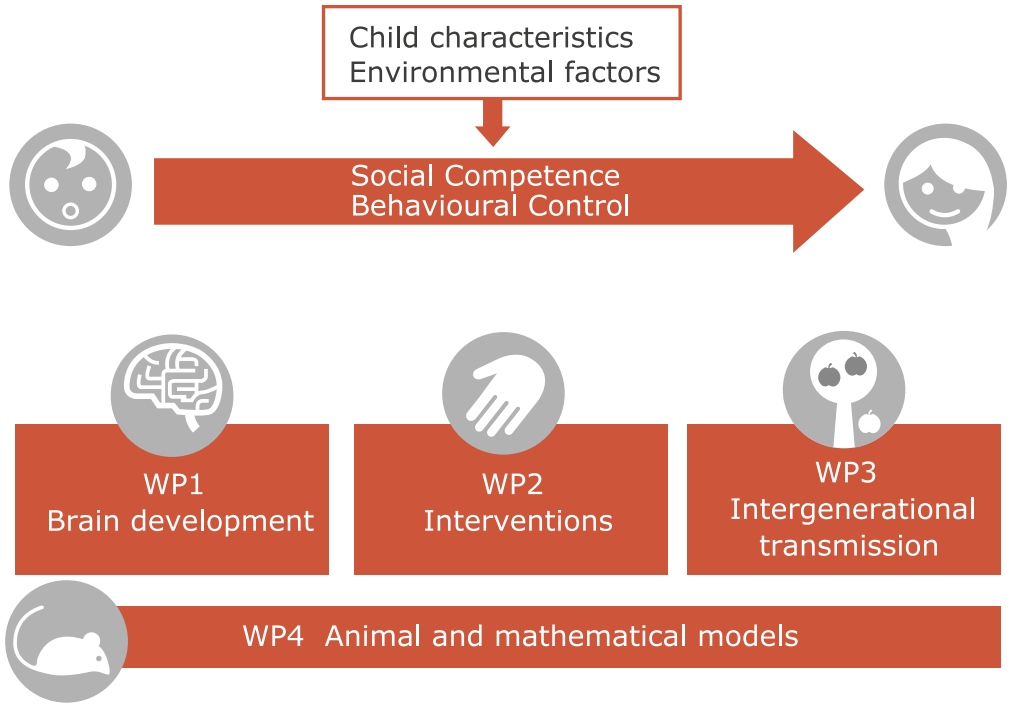


Figure 1. *The CID consortium with its four work packages.*

The studies reported in this thesis are part of WP3 of this CID consortium, which is about intergenerational transmission. Intergenerational transmission refers to the genetic and environmental transmission of characteristics, skills and traits from parents to their children and creates for example resemblance between siblings. In what follows I will introduce the two topics of this thesis, bullying and parental age at birth of offspring, and the studies within that topic. Bullying is an outcome trait for which I looked at risk factors and familial resemblance. Parental age at birth of offspring is a predictor variable that represents a characteristic of parents which is thought to influence multiple aspects of children's development.

1. Bullying

Bullying is a widespread phenomenon. It is the most frequent form of abuse encountered by children. Around 40% of the children are victims of bullying; around 30% of the children are perpetrators of bullying (Modecki, Minchin, Harbaugh, Guerra & Runions, 2014) and it is thought that being a victim and being a perpetrator may be associated within the same individual. Bullying has detrimental consequences on both children's physiological and psychological health. Victims have, for instance, more often internalizing problems and anxiety, or depression disorder, while bullies are more at risk for delinquent behavior (Wolke & Lereya, 2015).
These negative health effects persist into adolescence and adulthood. That is, being victimized in school is a risk factor for the manifestation of depression later in life (Wolke & Lereya, 2015). In my PhD I investigated children's vulnerability of bullying in schools, employing a large sample of twins of the NTR.

1.1 Risk Factors

Factors influencing involvement in bullying range from individual to contextual. Studies so far mostly investigated general factors, that are not specific to twins. These studies showed that boys are more vulnerable to become victims or bullies (e.g., Iossi Silva, Pereira, Mendonça, Nunes & Oliveira, 2013), but they are inconsistent about the effect of age (Moura, Cruz, Quevedo, 2011; Craig et al., 2009). Questions regarding bullying in twins (i.e., twin-specific factors) remain unanswered. For instance, it is unknown what the effects are of having a co-twin (i.e., having close companionship) or sharing a classroom with a close sibling such as a cotwin. The first question is about whether twins - having a co-twin by their side - are at higher, equal, or lower risk than non-twin children (called singletons). So far three studies tried to answer this question for victimization by comparing twins with unrelated singletons (Barness & Boutwell, 2013; Oshima et al., 2010; Weissenberg et al., 2007). Studies should however compare twins with their non-twin siblings to make sure they closely match on important family background factors. The second question is about whether or not twins should share classroom in primary school. This is an important question, also in the light of the policy of many schools to separate twins and not allow them to be in the same classroom.

Chapter 2 reports on a study that investigated the prevalence of the active form of bullying ("perpetration") and the passive form ("victimization"). First, in this chapter I compared the bullying behavior of twins with that of singletons to see, for instance, whether having a cotwin protects children from being bullied. We are the first that will do so with a large sample of twins and their non-twin siblings in order to make sure they closely match on important family-background factors. Importantly, we were the first to compare twins and singletons on the prevalence of perpetration.

Secondly, this chapter describes how the prevalence of perpetration and victimization is influenced by twin-specific (e.g., classroom-sharing) and non-twin specific (e.g., age and gender) factors. Whether or not twins should share a classroom is an important decision for parents of twins who enter primary school. When we know the effect of classroom sharing on bullying, we can give twin parents and teachers advice regarding the placement of twins, since classroom-sharing is a malleable factor.

1.2 Etiology

Even after accounting for general effects, like boys being more involved in bullying than girls, large individual differences remain. So it is not the case that all boys and none of the girls are bullies. Why are some children more likely to be involved in bullying, either as a victim, bully, or both (i.e., bully-victim)? The literature does not include many studies that investigated causes of individual differences.

The genetically-sensitive studies done so far studied mostly victimization (e.g., Silberg et al., 2016). Perpetration, and the association between both, was investigated only once (Ball et al., 2008). Furthermore, almost all studies are based on bullying behavior *in general*, while we know from earlier studies that bullying behavior comes in different forms, and that each have their own prevalence. Boys and girls are differently involved in these forms: boys are more often involved in physical and verbal bullying, while girls are more often involved in relational bullying (Scheithauer, Hayer, Petermann & Jugert, 2006). Eastman et al. (2018) were the first that investigated the genetic and environmental influences on different forms of victimization. However, this remains unanswered for different types of perpetration and for the overlap between them.

Chapter 3 is about the relative contributions of genetic and environmental influences on different forms of bullying. It is about the question why some children are involved in bullying, either as victims or as perpetrators or even both. Genetic and environmental effects were investigated for general, verbal, physical, and relational bullying, for both the active form (perpetration) and the passive form (victimization). I also looked at the question of their association (representing bully-victims) across these different forms. Information on bullying was obtained from children's teachers in primary school. The parents in families registered with the Netherlands Twin Register were asked for permission to approach their children's teachers and if they agreed, the teachers were asked to complete a survey including questions about bullying.

2. Parental Age

The influences of advanced parenthood on various severe neurodevelopmental disorders are well established. That is, offspring of older parents are at higher risk for schizophrenia, autism, and Down Syndrome (Merikangas, 2016; 2017; de Kluiver et al. 2017). Consequently, concerns are growing, because the mean age at which people start a family has been growing since the widespread availability of contraception around 1970. While in the Netherlands in 1970 the mean age at first birth was 24.3, nowadays the mean age at first birth is around 30 (CBS, 2019). The important next step in this field is to investigate whether the established adverse effects of advanced parental age on rare disorders extent to milder – but more common – neurodevelopmental problems. This part of this thesis addresses this question, regarding 1) two broad categories of behavioral and emotional problems in children, i.e. internalizing and externalizing problems, and 2) attention problems and cognitive functioning as assessed by psychometric IQ tests and educational attainment (CITO scores). Both projects are unique in that they included and combined data of all four large CID cohorts of WP3, representing a very large sample of children from all across the Netherlands.

The cohorts that are part of CID WP3 are Generation-R (Gen-R), Tracking Adolescents' Individual Lives Survey (TRAILS), Research on Adolescent Development and Relationships-Young (RADAR-Y), and Netherlands Twin Registry (NTR). Gen-R is conducted by Erasmus University Rotterdam, TRAILS by the University of Groningen, RADAR-Y by Utrecht University, and NTR by the Vrije Universiteit Amsterdam. Gen-R is a prospective study that recruited mothers in the city of Rotterdam during pregnancy. Their partners and their children were also invited to participate. TRAILS recruited their sample in the Northern regions of the Netherlands and RADAR-Y in the province of Utrecht and four large cities in the mid-west part of the Netherlands. NTR recruits new-born twins form all regions in the Netherlands. These CID cohorts follow children in longitudinal studies.

2.1 Internalizing and Externalizing Problems

There is little comprehensive evidence from population-based cohorts on the effect of paternal and maternal age on internalizing and externalizing problem behavior. A number of studies have investigated effects on externalizing problems. While most studies have identified negative associations with maternal age (for a review, see: Tearne, 2015), there are exceptions. For instance, Weiser et al. (2008) found that advanced parental age is associated with poorer social functioning. These prior studies, however, used a wide variety of populations and analytic strategies. In addition, the effects are mostly investigated for mothers' age, but might be different for fathers' age, underlining the need to study them separately.

Saha, Barnett, Buka, McGrath (2009) indeed showed different effects for maternal and paternal age. They showed that advanced maternal age was protective against externalizing problems (but associated with an increased risk for internalizing problems), while advanced paternal age was associated with an increase of adverse externalizing problems (and not with internalizing problems). Importantly, prior population-based studies rarely included internalizing problems. There are a few exceptions, but these exceptions require replication in other cohorts (e.g., Orlebeke, Knol, Boomsma & Verhulst, 1998; Tearne, Robinson, Jacoby, Newnham, McLean, 2015). Taken together, there is need for studies that investigate both the paternal and maternal age effects on core dimensions of offspring mental health, including both internalizing and externalizing problems, with the same analytical strategy.

Chapter 4 is about the parental age effects on externalizing and internalizing problem behavior. Externalizing and internalizing data of participants around age 10-13 years in four population-based cohorts of CID were analyzed. The four cohorts have reports from mothers, fathers, teachers, and the children themselves. As the perception of childhood problems may differ for these different informants for example because of situation-specific problem behavior, we used a multiple informant design in order to get a comprehensive set of outcome measures. The reports from teachers are particularly valuable, because their reports are unlikely to be affected by parental age-related report biases. We tested both linear and nonlinear effects, with and without adjusting for child gender and socio-economic status. Due to the previous mixed findings and the availability of a very large sample of children, we used a cross-validation approach. That is, we generated hypotheses based on one random half of the data of each cohort and evaluated these set of hypotheses within the other half of the data. The large datasets of the cohorts were thus of great value. A Bayesian statistical framework was used to investigate the overall support (i.e., of all cohorts together) for each hypothesis about the possible risk of young and older parenthood.

2.2 Attention Problems & Cognitive Functioning

Apart from effects of parental age on internalizing and externalizing problems, we investigated the effects on attention problems and cognitive functioning, which are inversely associated. ADHD is one of the most common neurodevelopmental disorders in childhood (Faraone, Sergeant, Gillberg & Biederman, 2003). All children with ADHD suffer from attention problems. Most studies showed that offspring of *younger* mothers are more at risk (e.g., Mikkelsen et al., 2016), but this effect is also found for *older* mothers (e.g., Chudal et al., 2015). Mikkelsen et al. (2016) found no effect for fathers, while D'Onofrio et al. (2014) reported a higher risk for *older* fathers and Chudal et al. (2015) for *younger* fathers. The mixed findings for ADHD are mirrored in the literature on effects on offspring's cognitive functioning. For example, D'Onfrio et al. (2014) showed adverse effects of delayed parenthood, while McGrath et al. (2013) showed adverse effects of young fatherhood as well. The previous mixed findings for ADHD and cognitive functioning might be due to differences between cohorts and analytic strategies. Hence, in this study we combined data of four large cohorts and used the same analytic procedure for each cohort. Again, the dataset was split into a discovery and a replication sample. The age effects for fathers and mothers might differ and might be the reverse from what we would have expected from research about the more severe neurodevelopmental outcomes.

Chapter 5 looks into the effects of parental age on the neurodevelopmental outcomes attention problems and cognitive functioning. For cognitive functioning, we focused on intelligence (psychometric IQ) and educational achievement (as measured by the Dutch CITO test which is completed by the majority of children in the Netherlands). The aim was to clarify whether effects for relatively rare disorders extend to these more common outcomes (i.e., the entire distribution). Again, data of mothers, fathers, teachers and the children themselves were used from the same four large CID cohorts. Paternal and maternal age were separately investigated, both with and without the adjustment of child gender and socio-economic status of the parents. In this chapter, the same analytical method was used as in Chapter 4 to combine the data of the four cohorts in order to investigate the possible risks of young and older parenthood on attention problems, academic achievement and IQ.

A summary of the main results of all four projects and a general discussion can be found in **Chapter 6**, followed by a Dutch summary in **Chapter 7**.

References

Centraal Bureau voor de Statistiek. (2019, January 23). Vrouwen steeds later moeder. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2018/05/vrouwen-steeds-later-moeder

Chudal, R., Joelsson, P., Gyllenberg, D., Lehti, V., Leivonen, S., Hinkka-Yli-Salomäki, S., ... & Sourander, A. (2015). Parental age and the risk of attention-deficit/hyperactivity disorder: a nationwide, population-based cohort study. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY, 54(6), 487-494.*

Craig, W., Harel-Fisch, Y., Fogel-Grinvald, H., Dostaler, S., Hetland, J., Simons-Morton, B., ... & Pickett, W. (2009). A cross-national profile of bullying and victimization among adolescents in 40 countries. *INTERNATIONAL JOURNAL OF PUBLIC HEALTH, 54(2), 216-224.*

De Kluiver, H., Buizer Voskamp, J. E., Dolan, C. V., & Boomsma, D. I. (2017). Paternal age and psychiatric disorders: A review. *AMERICAN JOURNAL OF MEDICAL GENETICS PART B: NEUROPSYCHIATRIC GENETICS, 174(3), 202-213.*

D'onofrio, B. M., Rickert, M. E., Frans, E., Kuja-Halkola, R., Almqvist, C., Sjölander, A., ... & Lichtenstein, P. (2014). *PATERNAL AGE AT CHILDBEARING AND OFFSPRING PSYCHIATRIC AND ACADEMIC MORBIDITY. JAMA PSYCHIATRY, 71(4), 432-438.*

Eastman, M. L., Verhulst, B., Rappaport, L. M., Dirks, M., Sawyers, C., Pine, D. S., ... & Roberson-Nay, R. (2018). Age-Related Differences in the Structure of Genetic and Environmental Contributions to Types of Peer Victimization. *BEHAVIOR GENETICS,* 48(6), 421-431.

Faraone, S. V., Sergeant, J., Gillberg, C., & Biederman, J. (2003). The worldwide prevalence of ADHD: is it an American condition?. *WORLD PSYCHIATRY, 2(2), 104.*

Houtrow, A. J., Larson, K., Olson, L. M., Newacheck, P. W., & Halfon, N. (2014). Changing trends of childhood disability, 2001–2011. *PEDIATRICS, 134(3), 530-538.*

Hvolgaard Mikkelsen, S., Olsen, J., Bech, B. H., & Obel, C. (2016). Parental age and attention-deficit/hyperactivity disorder (ADHD). *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY, 46(2), 409-420.*

Iossi Silva, M. A., Pereira, B., Mendonça, D., Nunes, B., & Oliveira, W. A. D. (2013). The involvement of girls and boys with bullying: an analysis of gender differences. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH, 10(12), 6820-6831.*

McGrath, J., Mortensen, P. B., Pedersen, C. B., Ehrenstein, V., & Petersen, L. (2013). Paternal age and general cognitive ability—a cross sectional study of Danish male conscripts. *PLOS ONE, 8(10), E77444.*

Merikangas, A. K., Calkins, M. E., Bilker, W. B., Moore, T. M., Gur, R. C., & Gur, R. E. (2017). Parental age and offspring psychopathology in the Philadelphia Neurodevelopmental Cohort. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY, 56(5), 391-400.*

Merikangas, A. K., Segurado, R., Kelleher, E., Hogan, D., Delaney, C., Gill, M., ... & Heron, E. A. (2016). Parental age, birth order and neurodevelopmental disorders. *MOLECULAR PSYCHIATRY, 21(6), 728.*

Modecki, K. L., Minchin, J., Harbaugh, A. G., Guerra, N. G., & Runions, K. C. (2014). Bullying prevalence across contexts: A meta-analysis measuring cyber and traditional bullying. *JOURNAL OF ADOLESCENT HEALTH, 55(5), 602-611.*

Moura, D. R. D., Cruz, A. C. N., & Quevedo, L. D. Á. (2011). Prevalence and characteristics of school age bullying victims. *JORNAL DE PEDIATRIA, 87(1), 19-23.*

Orlebeke, J. F., Knol, D. L., Boomsma, D. I., & Verhulst, F. C. (1998). Frequency of parental report of problem behavior in children decreases with increasing maternal age at delivery. *PSYCHOLOGICAL REPORTS, 82, 395-404. DOI: 10.2466/PR0.1998.82.2.395*

Saha, S., Barnett, A. G., Buka, S. L., & McGrath, J. J. (2009). Maternal age and paternal age are associated with distinct childhood behavioural outcomes in a general population birth cohort. *SCHIZOPHRENIA RESEARCH, 115(2-3), 130-135.*

Scheithauer, H., Hayer, T., Petermann, F., & Jugert, G. (2006). Physical, verbal, and relational forms of bullying among German students: Age trends, gender differences, and correlates. *AGGRESSIVE BEHAVIOR: OFFICIAL JOURNAL OF THE INTERNATIONAL SOCIETY FOR RESEARCH ON AGGRESSION, 32(3), 261-275.*

Tearne, J. E. (2015). Older maternal age and child behavioral and cognitive outcomes: a review of the literature. *FERTILITY AND STERILITY, 103(6), 1381-1391*

Tearne, J. E., Robinson, M., Jacoby, P., Li, J., Newnham, J., & McLean, N. (2015). Does late childbearing increase the risk for behavioural problems in children? A longitudinal cohort study. *PAEDIATRIC AND PERINATAL EPIDEMIOLOGY, 29, 41-49. DOI: 10.1111/PPE.12165.*

Weiser, M., Reichenberg, A., Werbeloff, N., Kleinhaus, K., Lubin, G., Shmushkevitch, M., ... & Davidson, M. (2008). Advanced parental age at birth is associated with poorer social functioning in adolescent males: shedding light on a core symptom of schizophrenia and autism. *SCHIZOPHRENIA BULLETIN, 34(6), 1042-1046.*

Wolke, D., & Lereya, S. T. (2015). Long-term effects of bullying. *ARCHIVES OF DISEASE IN CHILDHOOD, 100(9), 879-885.*

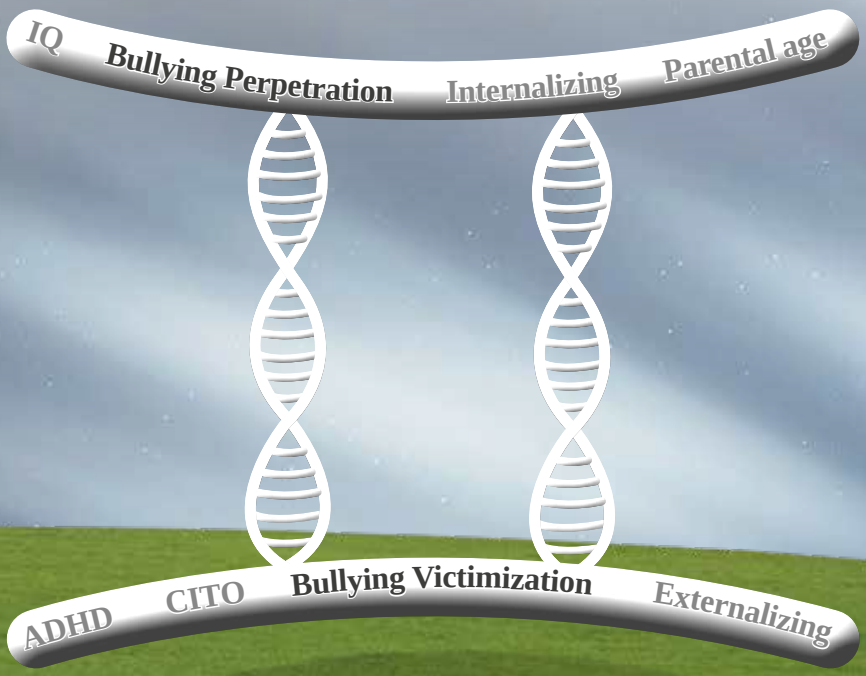
Part 1

Bullying

Chapter 2

Bullying and Victimization:

the Effect of Close Companionship



This chapter is based on:

Veldkamp, S. A.M., van Bergen, E., de Zeeuw, E. L., Van Beijsterveldt, C. E., Boomsma, D. I., & Bartels, M. (2017). Bullying and victimization: The effect of close companionship. ***Twin Research and Human Genetics, 20*** **(1), 19-27.**

Abstract

Peer bullying and victimization is a widespread phenomenon among school-age children and can have detrimental effects on the development of children. To examine whether having a close companion during childhood increases or decreases risk of victimization and bullying this study compared twins to singleton children. A large group of twins (*N*=9,909) were included who were compared to their related non-twin siblings (*N*=1,534) aged 7-12 from the Netherlands Twin Register, thus creating optimal matching between twins and non-twins. Bullying and victimization were each based on a four-item scale filled out by their teachers. Prevalence rates for either bullying or victimization did not differ between twins and singletons. In total, in the past couple of months 36% of children bullied peers moderately to severely, and 35% suffered moderately to severely from victimization. Boys were more likely to bully and were more prone to become a victim than girls. The most notable finding is that female twin pairs placed together in the same classroom did not bully more often, but were victimized less often, thus pointing to a protective effect of having a close companion in the classroom.

Introduction

Peer bullying, a widespread phenomenon among school-age children, can have detrimental consequences on childhood development. Both bullies and victims of bullying suffer more from a variety of problems than their uninvolved peers, including psychosomatic problems (Gini & Pozzoli, 2009), low self-esteem, anxiety, depression, loneliness, and low social self-concept (Hawker & Boulton, 2000). Consequently, this topic has attracted a lot of attention from educational practitioners, mental health services, and academics. This study compares the prevalence rates of bullying and victimization in twins and singletons to learn something about the potentially protective effects of having a close peer companion during childhood.

It still remains an open question whether children with a close peer companion, e.g., twins, are at higher, lower, or similar risk for bullying and victimization. One might, on the one hand, think that these strong companionship dyads are more vulnerable to be victimized than singletons. For example, Hay and Preedy (2006) suggest that the strong relationship in twins might restrict the interaction with other children. Research indeed showed that twins show less prosocial behavior than singletons when they play with an unfamiliar peer (DiLalla, 2006), which could make twins more often the target of peer victimization. On the other hand, it is well known that the presence of friends protect children from being bullied (Hodges, Boivin, Vitaro & Bukowski, 1999; Goldbaum, Craig, Pepler & Connolly, 2003) and it seems that sibling relationships provide protection against peer bullying as well (Lamarche et al., 2006).

To our knowledge, only three studies have been carried out to detect twin-singleton differences and all three only looked at victimization (Barnes & Boutwell, 2013; Oshima et al., 2010; Weissenberg, Landau & Madgar, 2007). The results of these studies are contradictory. Barnes and Boutwell (2013) found that singletons are at higher risk, while Weissenberg, Landau and Madgar (2007) found the opposite effect, and Oshima et al. (2010) reported no effect at all. It should be noted that all these studies were based on unrelated singletons, meaning that the twin and singleton samples may differ on important family-background factors, which hampers twin-singleton comparisons. Furthermore, the latter two studies suffered from low power (N = 341 and N = 72, respectively). In contrast, we employed a large sample ($N \approx$ 10,000) of twins and their non-twin siblings, making them closely matched on family background. Importantly, the three previous studies solely focused on victimization. Our study is the first to examine twin-singleton differences in bullying as well.

For bullying, we might speculate that twins are more involved than singletons. Twins have a pal with whom they could bully together, which is not the case for singletons.

Questions also remain about whether bullying in twins depends on twin-specific characteristics, such as the pair's zygosity and gender composition. Studies so far have shown no effect of zygosity on the proportion of bullies and victims in twins (e.g. Ball et al., 2008; Brendgen et al., 2008; Lamarche et al., 2007; Shakoor et al., 2015). However, it is still unknown whether the prevalence of bullying is affected by the gender composition of the pair (same-gender or mixed gender) and whether this effect changes as children age. Given that children increasingly play with same-sex peers (Rose & Rudolph, 2006), it might be hypothesized that the effect of gender composition is absent in the beginning and is present at the end of primary school.

All the potential twin-specific influences discussed so far are not malleable, but whether a twin pair attends the same or different classrooms is usually a choice. Hence, an important question for education is whether the risk of bullying in twins is related to whether or not they attend the same classroom, and whether that effect differs for monozygotic (MZ) and dizygotic (DZ) twins. The previously mentioned effects for twins may be more pronounced for twin pairs attending same classrooms. Twins in the same classroom always have their close relative by their side who can give them support and can protect them from being bullied. Classroom-sharing might therefore have a protective effect regarding victimization. We hypothesize that this effect might be stronger for MZ twins, since they tend to be more attached to each other than are DZ twins (Tancredy & Fraley, 2006). There is one relatively small study so far that has addressed this classroom-sharing question, suggesting that class-sharing might protect against victimization (Lamarche et al., 2006). In the current study, this previous finding was tested in a 20 times larger sample to overcome power issues. Classroom sharing might also affect the proportion of bullies, because twins in same classrooms can bully classmates together. To our knowledge, previous research has not addressed this issue. This study, therefore, tested classroom-sharing effects in bullying as well.

In previous research, some non-twin specific characteristics have been tested, namely gender and age. Most studies show that boys are more likely to bully than girls (e.g. Bowes et al., 2013; Kokkinos & Antoniadou, 2013; Von Marées & Petermann, 2010). An explanation why boys are more prone to be bullies than girls, is that they are generally more aggressive. Aggression is the overarching concept of which bullying is a subset (Griffin & Gross, 2004).

For victimization findings are less consistent. There is also no agreement about the trend of bullying and victimization during primary school. Some studies show that the prevalence rate of bullying rises as children age (Atik & Güneri, 2013), while others show the opposite or no clear effects at all (e.g. Camodeca, Goossens, Terwogt & Schuengel, 2002; Pellegrini & Long, 2002).

Firstly, the current study investigated whether twins are at higher, lower, or similar risk compared to their non-twin siblings for both bullying and victimization. Secondly, this study tested whether the risk for bullying and victimization differs for MZ versus DZ twins, same-sex versus opposite-sex twins, and twins attending the same versus separate classes. Meanwhile, age and gender effects were investigated. Thirdly, it explored whether a possible classroom effect differs for MZ and DZ twins and whether the effect of gender composition changes over time.

Method

Participants

The current study used participants of the Netherlands Twin Register (NTR; Beijsterveldt et al., 2013). The NTR is established by the Department of Biological Psychology at the Vrije Universiteit Amsterdam. The reported project was approved by the medical ethical committee of the Vrije Universiteit Amsterdam (NTR/25-05-2007). Parents of the twins aged 7, 9/10 and 12 are asked for their consent to approach the teachers of the twins and their non-twin siblings with a survey. The survey for the primary school teachers includes items on bullying and victimization since 2010. The current study includes data collected between 2010 and 2015. A subset of the final sample of twins and singletons had data on two (twins: *N*=1,579 individuals; siblings: *N*=162) or three (twins: *N*=92 individuals; siblings: *N*=1) time points due to the longitudinal data collection protocol. Multiple time points were included as multiple cases, while statistically controlling for their non-independence (see section Statistical Analyses).

The twins and siblings were born between 1997 and 2008. The following figures refer to number of data points, not individuals. Surveys of twins were excluded if twin zygosity was unknown (*N*=193), if they were filled out by someone other than the regular teacher (*N*=81), if familiarity with the student was below average (*N*=74), if the survey was filled out by the same teacher while the twins were in separate classes (*N*=11), or if twin-pairs attending the same class were rated by different teachers (*N*=108). This resulted in a total twin-sample of 10,063 cases for the calculation of bullying and victimization scores. This sample included mostly cases for twin pairs for whom data was available for both twins (*N*=4,337 pairs).

Incomplete data (*N*=1,389) were mostly due to only one of the teachers returning the survey (when twins were in separate classes, *N*=1,230). In addition, sibling data (*N*=1,534) were included in order to compare the prevalence rates of bullying and victimization for twins and singletons. Subsequently, surveys were excluded when there were more than two missing values on the bullying and victimization scale, described in the measurement section (*N*=167; *N*=154, respectively). This resulted in a total sample of 9,896 cases for bullying and 9,909 cases for victimization. Only a few twins (*N*=73) and siblings (*N*=10) had a missing value on one of the traits. The age of the twins ranged from 6.52 to 12.94 years (*M*=9.49, *SD*=2.01), for the siblings this was 4.68-13.43 (*M*=9.99, *SD*=1.67).

Measures

Bullying and victimization were measured by four items each, scored on a five-point scale: from 0 (*never*), 1 (*once or twice*), 2 (*two or three times a month*), 3 (*about once a week*) to 4 (*several times a week*). The items for victimization assessed 1) "how often has the child been victimized in the past couple of months? (in general)", 2) "how often has the child been teased, laughed at, or called names in the past couple of months? (verbal victimization)"; 3) "how often has the child been physically victimized, such as being hit, kicked and pushed in the past couple of months? (physical victimization)"; and 4) "how often has the child been excluded by other children, ignored, or have other students spread false rumors? (relational victimization)". Bullying was assessed with the same items, but in the perpetration form (see Jansen et al., 2012). The total score of both phenotypes could range from 0 to 16.

The reliability of the questions was good for both bullying (α = .84) and victimization (α = .80). Sum scores were computed when there was at most one missing item for a scale. Other missing items were imputed by the rounded averaged-item score of the scale for that child. The sum scores had an "L"-shaped distribution, which was divided into three categories to determine the prevalence rates of bullying and victimization. The first category for both bullying and victimization was defined by a sum score of 0 and was labeled as "never bullied/was never victimized". The second category was defined by a sum score of 1 or 2, ranging from (very) mild to moderately bullied/victimized and was labeled as "moderately bullied/was moderately victimized". The last category was defined by a score of at least 3 and ranged from substantial to very severe, labeled as "severely bullied/was severely victimized". A score of 3 means that someone scored "once a week" on one of the four items. In subsequent statistical analyses, raw item data was used to avoid bias in parameter estimates due to non-normality of the sum scores. The four items for bullying and victimization were combined into factor-scores for each child.

Simultaneously, the factorial level of the model was used to explain the effects of our predictors on bullying and victimization.

Statistical Analyses

Data were prepared in the statistical software R, version 3.2.0 (R core team, 2015) and analyzed in Mplus version 6 (Muthén & Muthén, 1998-2012). To accommodate multiple testing, we used adjusted p-values according to the Holm-Bonferroni method (Holm, 1979). This method is a less strict correction for multiple testing than the Bonferroni method. It is a sequential procedure that uses a less strict rejection criterion for each subsequent comparison to reduce the Type II (false negative) errors. Type II errors arise from decreasing Type I errors in multiple testing.

Twins and siblings came from the same family and thus were well-matched for confounders. To correct for non-independent observations, which results in underestimated standard errors when not taken into account, the analyses were corrected by using tests based on the sandwich or Huber/White variance estimator (Williams, 2000). The subset of children that had data on multiple time points was also be used in the analyses, since the sandwich variance estimator also corrected for this dependency.

The statistical analyses included descriptive statistics on prevalence rates for the total sample, as well as for twins versus singletons, boys versus girls, MZ twins versus DZ twins, and twins attending same versus separate classrooms. Subsequently, main and interaction effects were statistically tested.

Main analyses

To test for differences in prevalence rates for the different subgroups, a regression was performed. First, it was tested whether twins are at higher, lower or similar risk compared to their non-twin siblings for both bullying and victimization.

Second, it was tested whether the risk for bullying and victimization differs for MZ and DZ twins, same-sex versus opposite sex-twins, and whether the risk is lower if twins attend the same or separate classrooms. At the same time, age and gender effects were investigated.

The second model included the main effects of age (both linear and curvilinear), classroom sharing, gender, and zygosity.

One variable that indicates whether the twins are of same- or opposite-sex was added as a covariate to correct for confounding gender effects in the zygosity variable, as DZ pairs can and MZ pairs cannot be of mixed genders. The age predictor variables in the model were entered as mean-centered continuous variables, while the other variables were entered as dichotomous variables.

In previous research, it has been shown that classroom assignment is associated with Social Economic Status (SES) and externalizing problems (van Leeuwen, van den Berg, van Beijsterveldt & Boomsma., 2005). If an effect of classroom sharing were to be found, the model would be extended with the possible confounders SES and externalizing- and internalizing problems to see whether the outcome changes. Data on these variables at age 3, i.e. before children went to school, are present for two-thirds of our full sample (see van Leeuwen et al. (2005) for measurement details).

Interaction Analyses

If a classroom-sharing effect were to be found, a simplified model would be fitted to test whether the effect of classroom sharing differs for being a girl-girl, boy-boy, or opposite-sex twins. To test this interaction effect, the covariates gender composition (same-sex versus opposite-sex), gender, and zygosity were deleted from the main model to test this interaction effect fairly. In addition, the variable that specifies the type of twin pair was added as two dummy variables in the two interaction terms with classroom placement. The first dummy variable was coded 1 for girl-girl twins, and the second coded 1 for opposite-sex twins, leaving the boy-boy twin pairs to represent the reference group.

In the follow-up analyses, we first examined whether gender composition (same-sex versus opposite-sex) interacts with age in predicting bullying and victimization. By adding this interaction term, it was tested whether the effect of being a same-sex versus opposite-sex twin on bullying and victimization changes as children age. In addition, it was tested whether the effect of attending the same versus separate classrooms was different for MZ and DZ twins. The main effects were still included in the model to correct for possible confounding effects.

Results

Descriptive statistics

Table 1 presents the prevalence rates of bullying and victimization for the total sample, for twins versus singletons, boys versus girls, MZ versus DZ twins, and twins in same versus separate classes. As can be seen, 36.2% of the total sample (including twins and siblings) has bullied others, while 34.8% has been victimized in the past couple of months.

Table 1. *Prevalence Rates of Bullying and Victimization*

	Bullying				Victimization			
	N	**Never (%)**	**Moderate (%)**	**Severe (%)**	***N***	**Never (%)**	**Moderate (%)**	**Severe (%)**
Total sample	11,430	63.8	20.3	15.9	11,441	65.2	22.2	12.6
Twins vs. Sibs								
Twins	9,896	63.7	20.2	16.1	9,909	65.4	22.0	12.6
Singleton	1,534	64.1	21.1	14.8	1,532	63.6	23.3	13.1
Gender								
Boys	4,985	56.2	22.9	20.9	4,988	62.2	23.1	14.6
Girls	4,911	71.4	17.5	11.2	4,921	68.7	20.8	10.5
Zygosity								
MZ	3,725	65.2	19.3	15.5	3,735	67.3	20.9	11.8
DZ	6,171	62.8	20.7	16.4	6,174	64.3	22.7	13.0
Class								
Same class	5,111	64.5	19.7	15.8	5,116	68.1	20.6	11.3
Different class	4,677	62.8	20.8	16.4	4,684	62.2	23.7	14.1

Main analyses

Twins did not differ significantly from singletons with respect to bullying (β = -.015, *SE* = .017, p = .384) and victimization (β = .018, *SE* = .018, p = .313).

Table 2 shows that there was a significant main effect of gender indicating that boys were more likely to bully than girls. Also, a curvilinear age effect was found. The negative beta of the quadratic term implies that the curve for the age effect regarding bullying has an inverted U-shape. The raw data showed that bullying peaks around 9 years old. No other predictors approached significance.

Table 2. *Influences of the (Non) Twin-Specific Factors on Bullying (N=9,788)*[a]

Covariates	Coding		*β*	*SE*	*p*-value	*α-level* [b]
	0	1				
Gender	Male	Female	-.460	.030	<.001	.008
Zygosity	MZ	DZ	.020	.038	.601	
Class	Different	Same	-.030	.030	.315	
Same-sex	Same	Oposite	.055	.037	.139	.013 (*NS*)
Age	Continuous variable		.008	.008	.284	
Age squared	Continuous variable		-.018	.004	<.001	0.10

Note. SE = Standard Error. *OR* = Odds Ratio. *NS* = not significant.

[a] *N*=108 missing values on class-variable.

[b] Corrected α-level according to the Holm- Bonferroni method (Holm, 1979).

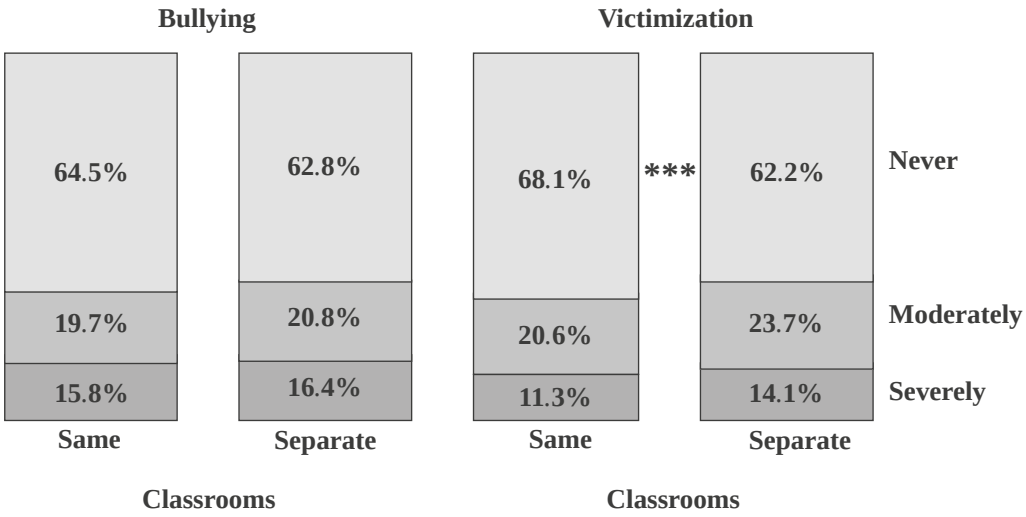


Figure 1. *Prevalence rates of bullying and victimization for twins attending same and different classes. Twins in the same classroom do not bully more often (left panel), but are bullied less often than those in separate classrooms (right panel). Follow-up analyses showed that this effect only holds for girl-girl twin pairs.*

Table 3 shows that for victimization both gender and classroom-sharing appeared to be significant predictors. Boys were more often victim than girls.

Table 3. *Influences of the (Non) Twin-Specific Factors on Victimizationg (N=9,800)*[a]

Covariates	Coding		*β*	*SE*	*p*-value	*α-level* [b]
	0	1				
Gender	Male	Female	-.236	.029	<.001	.008
Zygosity	MZ	DZ	.052	.037	.164	
Class	Different	Same	-.128	.030	<.001	.010
Same-sex	Same	Oposite	.069	.037	.059	.017 (*NS*)
Age	Continuous variable		.006	.008	.430	
Age squared	Continuous variable		-.028	.005	<.001	0.13

Note. SE = Standard Error. *OR* = Odds Ratio. *NS* = not significant.

[a] *N*=109 missing values on class-variable.

[b] Corrected α-level according to the Holm- Bonferroni method (Holm, 1979).

As for bullying, a curvilinear age effect was also found for victimization. The negative beta of the quadratic term implies that the curve for the age effect regarding victimization has an inverted U-shape. The raw data showed that the score peaked around age 9, so this pattern for victimization mirrors that of bullying. With respect to the twin-specific factors, the significant main effect of classroom sharing indicates that twins attending the same class were less often victim than those in separate classes. This effect is shown in Figure 1. Due to a possible non-random assignment of twins to same or separate classrooms, we redid the analysis taking SES and preexisting differences in externalizing-, and internalizing problems into account. After controlling for these possible confounders, classroom placement still had an effect of similar magnitude (β = -.138, *SE* = .036, $p < .001$). The other predictors, zygosity and gender composition, were not significant. In the follow-up analysis, we explored possible interaction effects.

Interaction results

The subsequent analysis showed that for boy-boy twins there is no effect of classroom sharing (β = -.002, *SE* = .044, p = .967). Compared to boy-boy twins, the effect of classroom sharing does not differ for opposite sex twins (β = -.062, *SE* = .055, p = .261), but does for girl-girl twins (β = -.335, *SE* = .057, $p < .001$), indicating that the classroom effect only holds for girl-girl twin pairs. This effect can also be seen in Table 4.

The second interaction term showed that the effect of gender decomposition (same versus opposite) does not change as children age with respect to bullying (β = -.019, *SE* = .016, *p* = .216) and victimization (β = -.011, *SE* = .016, *p* = .475). The third interaction term showed that the effect of classroom sharing does not differ for MZ and DZ twins for bullying (β = -.090, *SE* = .064, *p* = .159) and victimization (β = -.019, *SE* = .064, *p* = .772).

Table 4. *Victimization Rates for Girl-Girl, Boy-Boy and Opposite Sex Twin Pairs in Same versus Separate Classrooms*

		Victimization		
		Never (%)	**Moderate (%)**	**Severe (%)**
Girl-girl	Separate	66.5	22.8	10.7
	Same	73.8	18.1	8.1
Boy-boy	Separate	60.0	24.1	15.8
	Same	63.8	22.5	13.7
Oposite sex	Separate	60.1	24.2	15.7
	Same	66.5	21.4	12.1

Discussion

The goal of this study was to advance knowledge about protective and risk factors for bullying and victimization. By addressing important questions about twin-singleton differences and the influences of twin-specific characteristics we examined whether having a close companion during childhood increases or decreases risk of victimization and bullying. We report no twin-singleton difference, but do show that girl-girl twins are bullied less often when placed in the same classroom. Gender composition of the twin pair (same versus opposite sex), and twins' zygosity (MZ versus DZ), had no effect.

Our finding that twins are not at a higher or lower risk than singletons for bullying and victimization is in agreement with one of the three studies about twin-singleton differences regarding victimization (Oshima et al., 2010). Our finding, however, is inconsistent with the results of Weissenberg, Landau and Madgar (2007) and Barnes and Boutwell (2013), which showed higher risk and lower risk for twin respectively. The comparisons between the studies, though, is hampered by the use of unrelated singletons, which do not match on important family background factors. The first two studies (Oshima et al., 2010; Weissenberg, Landau & Madgar, 2007) had power problems as well.

Our study reduces both limitations by using large groups of twins and their non-twin siblings. Given the large sample size and the non-significant result, it can be concluded that no twin-singleton differences exist in prevalence rates regarding bullying and victimization. This finding supports the generalizability of twin studies regarding bullying and victimization to the non-twin population.

We, furthermore, showed that boys were significantly more likely to bully and to be victimized than girls and that the risk for both bullying and victimization peaks around 8-9 years. With respect to gender, this result is in line with the body of the literature on bullies (e.g. Bowes et al., 2013) and victims (e.g. Sentse, Kretschmer & Salmivalli, 2015; Takizawa, Maughan & Arseneault, 2014). For age, the existing literature was less clear. Some studies showed that the victimization decreases as children age (e.g. Sapouna, 2008), while others showed the opposite (Atik & Güneri, 2013). Bullying seems to be more stable. The current study, however, found a significant curvilinear age effect for both traits, showing that the risk for both traits peaks around 8-9 years. Although the effect of gender is large, the effect of age is rather small.

The risk for bullying and victimization is not influenced by twins' zygosity and the gender composition of the twin pairs. This is in line with previous studies that show no effect of twins' zygosity on involvement in bullying; either as bully or victim (e.g. Ball et al., 2008). For traits related to bullying, such as reactive and proactive aggression, also no zygosity differences exist (Lamarche et al., 2007). Regarding gender composition, Lamarche et al. (2006) found that 6-year-old twins of opposite sex are bullied more often. Our results, however, suggested that for both bullying and victimization same- versus opposite sex twins do not differ. Their study included 246 twin pairs, while we had the advantage of a nearly 20 times larger sample. In addition, they did not correct for multiple testing. For related phenotypes such as social independence, friendship, and behavior problems, no effect of gender composition has been found as well (Laffey-Ardley & Thorpe, 2006), which supports our results.

We hypothesized that the effect of gender composition is absent in the beginning and present at the end of primary school, given that children increasingly play with same-sex peers (Rose & Rudolph, 2006). We, however, found no interaction and concluded that the effect of gender composition does not change as children age.

The most notable finding of our study is that twin pairs that attend the same classroom do not bully more or less than separated pairs, but they seem to be less victimized.

This effect is restricted to girl-girl twin pairs, independent of zygosity, as it does not hold for boy-boy twin pairs or twin pairs of opposite sex. We know of only one study that investigated classroom effects regarding victimization (Lamarche et al., 2006). This study seemed to indicate that classroom sharing has a protective effect against victimization, however, their study and the effect size of the classroom effect were probably too small to reach significance. With our large sample, the difference in victimization scores did reach significance. The finding that girl-girl twins in the same classroom are bullied less often seems to indicate that classroom sharing has a protective effect for victimization. DiLalla and Mullineaux (2008) showed a protective effect of classroom sharing on peer problems as well; their peer-problems scale included one item (out of 4 items) about victimization. The protective effect of classroom sharing on victimization might be explained by, for instance, children's self-esteem. It is known that victims of bullying often suffer from low self-esteem (O'Moore & Kirkham, 2001). Girl twins in the same classroom might feel more self-confident by having a sister by their side, and are consequently less vulnerable for bullies. Another possible explanation might be the existence of sibling gender composition differences in the sibling relationship quality. The sibling relationship quality is best for girl-girl dyads (Buist, 2010), which might explain why the protective classroom effect is only present for girl-girl twin pairs. It is known that MZ twins are more likely than DZ twins to rate their co-twin as their best-friend (Foy, Vernon & Jang, 2001), but apparently this did not lead to a more pronounced classroom effect for MZ twins.

The result that girl-girl twins in the same classroom tend to be less victimized should be interpreted against the background of non-random classroom assignment. For example, the class assignment could have been influenced by preexisting cognitive, emotional, behavioral differences between the twins or other external factors. Indeed, van Leeuwen et al. (2005) showed that Dutch twin pairs from high SES families are more often separated. In addition, they showed that classroom assignment in the Netherlands is based on early childhood externalizing but not internalizing symptoms. However, after controlling for SES and preexisting differences in externalizing-, and internalizing problems at age 3, we showed that classroom sharing still had a protective effect of similar magnitude. This indicates that the protective effect of classroom sharing on victimization holds when taking factors linked to non-random classroom assignment into account.

To further investigate classroom assignment, we looked at data of a small subsample (N=66 twin pairs) that answered the question "Who decided to separate the twins?". The answer options "school", "parents" and "parents in agreement with the school" were each chosen by one-third of the sample.

This indicates that assignment of twins to same or separate classrooms is not systematic. We can conclude that the decision is at least not entirely based on child specific characteristics.

Notably, from all twin-specific factors discussed, classroom sharing is the only malleable factor. Classroom sharing is a malleable factor for schools that are large enough to have parallel year groups. Schools may have a set policy regarding classroom placement of twins (Saudino, Ronald & Plomin, 2005), depending on national context and beliefs of the principle and teachers. Nevertheless, most young twins and their parents prefer not to be separated (Gordon, 2015; Staton, Thorpe, Thompson & Danby, 2012). The protective effect of classroom sharing regarding victimization for girl-girl twins should be taken into account when twins enter primary school. We, furthermore, showed that twins in the same classroom do not bully more often. This is in line with our earlier findings that twins in the same classroom do not differ from twins in separate classrooms for other phenotypes as well, such as academic achievement, problem behavior (van Leeuwen et al., 2005; Polderman et al., 2010), and academic motivation (Kovas et al., 2015). Taken together, the placement of twins in the same classroom might be beneficial regarding victimization and is not harmful for other important behavioral outcomes.

While interpreting the results of our study it is important to acknowledge that bullying and victimization is based on teacher ratings and thus applies to the daily school setting. It is known that teachers do not necessarily rate bullying behavior in the same way as parents and children themselves do (Rønning et al., 2009), with agreement correlation in the range of .18 - .19 for bullying and .11 - .22 for victimization. Using teacher reported data is, however, not necessarily a disadvantage, since the view of teachers might be more objective.

Future research might investigate the extent to which twin separation at school entry is random. A true Randomized Control Trial (RCT) in which researchers decide whether a twin pair will be separated or not will face ethical resistance. We can, however, ask the teacher whether their school have a set policy regarding classroom assignment of twins. For the subset of schools that are large enough to have parallel year groups and that have set policies, whether a twin pair attends the same or different classes ought not to be related to child characteristics. Therefore, such a prospective study comes closer to a randomized trial. Follow-up research could also focus on the protective effect of these girl-girl twins attending same classrooms and try to uncover the underlying mechanisms. To speculate, this protective effect might be due to girl-girl twins feeling more self-confident when studying in the same classroom, which might not be the case for boy-boy and opposite-sex twins.

It is well known that more self-confidence makes children less prone to be a victim of peer bullying. Although further research is needed to elucidate the mechanisms, this finding in itself offers an important new perspective on the question if twin pairs should be separated or not.

To conclude, we demonstrate that assigning twins to the same classroom seems to have a protective effect for girl-girl twins, but not for boy-boy or opposite sex twins. Importantly, girl-girl twins in the same classroom do not bully more often. Our result indicates that it might be beneficial to keep girl-girl twins together when entering primary school.

References

Atik, G., & Güneri, O. Y. (2013). Bullying and victimization: Predictive role of individual, parental, and academic factors. *SCHOOL PSYCHOLOGY INTERNATIONAL,34(6), 658-673.*

Ball, H. A., Arseneault, L., Taylor, A., Maughan, B., Caspi, A., & Moffitt, T. E. (2008). Genetic and environmental influences on victims, bullies and bully victims in childhood. *JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY,49(1), 104-112.*

Barnes, J. C., & Boutwell, B. B. (2013). A demonstration of the generalizability of twin-based research on antisocial behavior. *BEHAVIOR GENETICS, 43(2), 120-131.*

Bowes, L., Maughan, B., Ball, H., Shakoor, S., Ouellet-Morin, I., Caspi, A., Moffitt, T.E., & Arseneault, L. (2013). Chronic bullying victimization across school transitions: the role of genetic and environmental influences. *DEVELOPMENT AND PSYCHOPATHOLOGY, 25(02), 333-346.*

Brendgen, M., Boivin, M., Vitaro, F., Dionne, G., Girard, A., & Pérusse, D. (2008). Gene-environment interactions between peer victimization and child aggression. *DEVELOPMENT AND PSYCHOPATHOLOGY, 20, 455–471.*

Buist, K. L. (2010). Sibling relationship quality and adolescent delinquency: A latent growth curve approach. *JOURNAL OF FAMILY PSYCHOLOGY, 24(4), 400.*

Camodeca, M., Goossens, F. A., Terwogt, M. M., & Schuengel, C. (2002). Bullying and victimization among school age children: Stability and links to proactive and reactiveaggression. *SOCIAL DEVELOPMENT, 11(3), 332-345.*

Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998). Revision and restandardization of the Conners Teacher Rating Scale (CTRS-R): factor structure, reliability, and criterion validity. *JOURNAL OF ABNORMAL CHILD PSYCHOLOGY,26(4), 279-291.*

Derks, E. M., Dolan, C. V., & Boomsma, D. I. (2004). Effects of censoring on parameter estimates and power in genetic modeling. *TWIN RESEARCH, 7(06), 659-669.*

DiLalla, L. F. (2006). Social development of twins. *TWIN RESEARCH AND HUMAN GENETICS, 9(01), 95-102.*

DiLalla, L. F., & Mullineaux, P. Y. (2008). The effect of classroom environment on problem behaviors: *A TWIN STUDY. JOURNAL OF SCHOOL PSYCHOLOGY, 46(2), 107-128.*

Foy, A. K., Vernon, P. A., and Jang, K. (2001). Examining the dimensions of intimacy in twin and peer relationships. *TWIN RESEARCH, 4:443-452.*

Gini, G., & Pozzoli, T. (2009). Association between bullying and psychosomatic problems: A meta-analysis. *PEDIATRICS, 123(3), 1059-1065.*

Goldbaum, S., Craig, W. M., Pepler, D., & Connolly, J. (2003). Developmental trajectories of victimization: Identifying risk and protective factors. *JOURNAL OF APPLIED SCHOOL PSYCHOLOGY, 19(2), 139-156.*

Gordon, L. M. (2015). Twins and kindergarten separation: Divergent beliefs of principals, teachers, parents, and twins. *EDUCATIONAL POLICY, 29(4), 583-616.*

Griffin, R. S., & Gross, A. M. (2004). Childhood bullying: Current empirical findings and future directions for research. *AGGRESSION AND VIOLENT BEHAVIOR, 9(4), 379-400.*

Hawker, D. S., & Boulton, M. J. (2000). Twenty years' research on peer victimization and psychosocial maladjustment: a meta-analytic review of cross-sectional studies. *JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY, 41(4), 441-455.*

Hay, D. A., & Preedy, P. (2006). Meeting the educational needs of multiple birth children. *EARLY HUMAN DEVELOPMENT, 82(6), 397-403.*

Hodges, E. V., Boivin, M., Vitaro, F., & Bukowski, W. M. (1999). The power of friendship: protection against an escalating cycle of peer victimization. *DEVELOPMENTAL PSYCHOLOGY, 35(1), 94.*

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *SCANDINAVIAN JOURNAL OF STATISTICS, 65-70.*

Jansen, P.W., Verlinden, M., Dommisse-Van Berkel, A., Mieloo, C., Van der Ende, J., Veenstra, R., Verhulst, F.C., Jansen, W., & Tiemeier, H. (2012). Prevalence of bullying and victimization among children in early elementary school: Do family and school neighborhood socioeconomic status matter? *BMC PUBLIC HEALTH, 12, 494.*

Kokkinos, C. M., & Antoniadou, N. (2013). Bullying and victimization experiences in elementary school students nominated by their teachers for Specific Learning Disabilities. *SCHOOL PSYCHOLOGY INTERNATIONAL, 34(6), 674-690.*

Kovas, Y., Garon-Carrier, G., Boivin, M., Petrill, S. A., Plomin, R., Malykh, S. B., Spinath, F., Murayama, K., Ando, J., Bogdanova, O.U., Brendgen, M., Dionne, G., Forget-Dubois, N., Galajinsky, E.V., Gottschling, J., Guay, F., Lemelin, J., Logan, J.A.R., Yamagata, S., Shikishima, C., Spinath, B., Thompson, L.A., Tikhomirova, T.N., Tosto, M.G., & Tremblay, R. (2015). Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *PERSONALITY AND INDIVIDUAL DIFFERENCES, 80, 51-63.*

Laffey-Ardley, S., & Thorpe, K. (2006). Being opposite: is there advantage for social competence and friendships in being an opposite-sex twin?. *TWIN RESEARCH AND HUMAN GENETICS, 9(01), 131-140.*

Lamarche, V., Brendgen, M., Boivin, M., Vitaro, F., Dionne, G., & Pérusse, D. (2007). Do friends' characteristics moderate the prospective links between peer victimization and reactive and proactive aggression?. *JOURNAL OF ABNORMAL CHILD PSYCHOLOGY, 35(4), 665- 680.*

Lamarche, V., Brendgen, M., Boivin, M., Vitaro, F., Pérusse, D., & Dionne, G. (2006). Do friendships and sibling relationships provide protection against peer victimization in a similar way? *SOCIAL DEVELOPMENT, 15(3), 373-393.*

Lynch, M., & Walsh, B. (1998). *GENETICS AND ANALYSIS OF QUANTITATIVE TRAITS (VOL. 1).* Sunderland, MA: Sinauer.

Muthén, L.K. and Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén

O'Moore, M., & Kirkham, C. (2001). Self esteem and its relationship to bullying behaviour. *AGGRESSIVE BEHAVIOR, 27(4), 269-283.*

Oshima, N., Nishida, A., Fukushima, M., Shimodera, S., Kasai, K., Okazaki, Y., & Sasaki, T. (2010). Psychotic-like experiences (PLEs) and mental health status in twin and singleton Japanese high school students. *EARLY INTERVENTION IN PSYCHIATRY, 4(3), 206-213.*

Pellegrini, A. D., & Long, J. D. (2002). A longitudinal study of bullying, dominance, and victimization during the transition from primary school through secondary school. *BRITISH JOURNAL OF DEVELOPMENTAL PSYCHOLOGY, 20(2), 259-280.*

Polderman, T. J., Bartels, M., Verhulst, F. C., Huizink, A. C., van Beijsterveldt, C. E., & Boomsma, D. I. (2010). No effect of classroom sharing on educational achievement in twins: a prospective, longitudinal cohort study. *JOURNAL OF EPIDEMIOLOGY AND COMMUNITY HEALTH, 64(01), 36-40.*

R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rønning, J. A., Sourander, A., Kumpulainen, K., Tamminen, T., Niemelä, S., Moilanen, I., Helenius, H., Piha, J., & Almqvist, F. (2009). Cross-informant agreement about bullying and victimization among eight-year-olds: whose information best predicts psychiatric caseness 10–15 years later?. *SOCIAL PSYCHIATRY AND PSYCHIATRIC EPIDEMIOLOGY, 44(1), 15-22.*

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: potential trade-offs for the emotional and behavioral development of girls and boys. *PSYCHOLOGICAL BULLETIN, 132(1), 98.*

Sapouna, M. (2008). Bullying in Greek primary and secondary schools. *SCHOOL PSYCHOLOGY INTERNATIONAL, 29(2), 199-213.*

Saudino, K. J., Ronald, A., & Plomin, R. (2005). The etiology of behavior problems in 7-year-old twins: substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. *JOURNAL OF ABNORMAL CHILD PSYCHOLOGY, 33(1), 113-130.*

Sentse, M., Kretschmer, T., & Salmivalli, C. (2015). The Longitudinal Interplay between Bullying, Victimization, and Social Status: Age-related and Gender Differences. *SOCIAL DEVELOPMENT, 24(3), 659-677.*

Shakoor, S., McGuire, P., Cardno, A. G., Freeman, D., Plomin, R., & Ronald, A. (2015). A shared genetic propensity underlies experiences of bullying victimization in late childhood and self-rated paranoid thinking in adolescence. *SCHIZOPHRENIA BULLETIN, 41(3), 754-763.*

Staton, S., Thorpe, K., Thompson, C., & Danby, S. (2012). To separate or not to separate? Parental decision-making regarding the separation of twins in the early years of schooling. *JOURNAL OF EARLY CHILDHOOD RESEARCH, 10(2), 196-208.*

Takizawa, R., Maughan, B., & Arseneault, L. (2014). Adult health outcomes of childhood bullying victimization: evidence from a five-decade longitudinal British birth cohort. *AMERICAN JOURNAL OF PSYCHIATRY.*

Tancredy, C. M., & Fraley, R. C. (2006). The nature of adult twin relationships: an attachment-theoretical perspective. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, 90(1), 78.*

van Beijsterveldt, C. E., Groen-Blokhuis, M., Hottenga, J. J., Franić, S., Hudziak, J. J., Lamb, D., Huppertz, C., de Zeeuw, E., Nivard, Schutte, N., Swagerman, S., Glasner, T., van Fulpen, M., Brouwer, C., Stroet, T., Nowothy, D., Ehli, E.A., Davies, G.E., Scheet, P., Orlebeke, J.F., Kan, K.J., Smit, D., Dolan, C.V., Middeldorp, C.M., de Geus, E.J., Bartels, M., & Boomsma, D.I. (2013). The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. *TWIN RESEARCH AND HUMAN GENETICS, 16(01), 252-267.*

Van Leeuwen, M., van den Berg, S. M., van Beijsterveldt, T. C., & Boomsma, D. I. (2005). Effects of twin separation in primary school. *TWIN RESEARCH AND HUMAN GENETICS, 8(04), 384-391.*

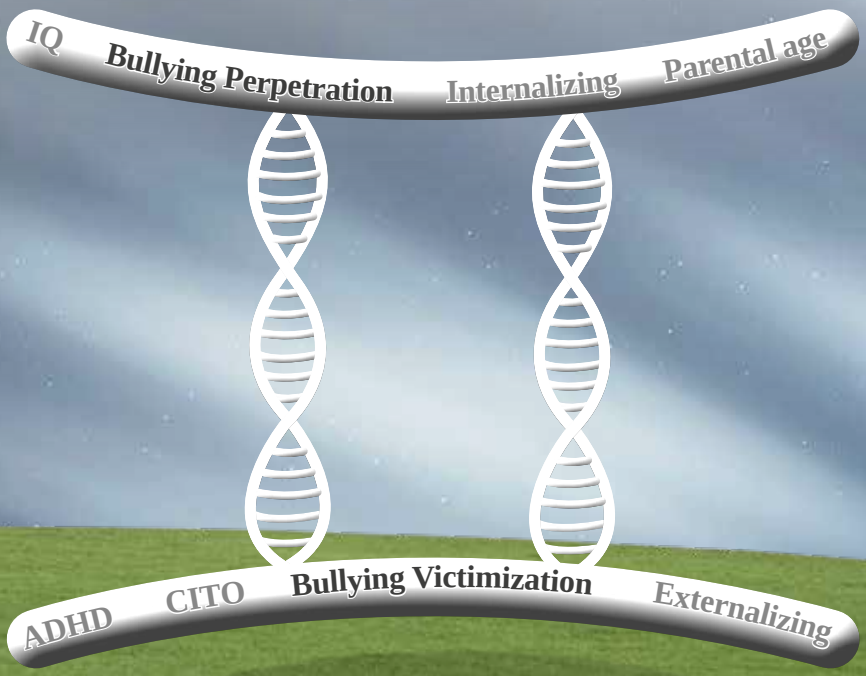
Von Marées, N., & Petermann, F. (2010). Bullying in German primary schools gender differences, age trends and influence of parents' migration and educational backgrounds. *SCHOOL PSYCHOLOGY INTERNATIONAL, 31(2), 178-198.*

Weissenberg, R., Landau, R., & Madgar, I. (2007). Older single mothers assisted by sperm donation and their children. *HUMAN REPRODUCTION, 22(10), 2784-2791.*

Williams, R. L. 2000. A note on robust variance estimation for cluster-correlated data. *BIOMETRICS 56: 645–646.*

Chapter 3

Genetic and Environmental Influences on Different Forms of Bullying Perpetration, Bullying Victimization, and their Co-occurrence



This chapter is based on:

Veldkamp, S. A.M, Boomsma, D. I., de Zeeuw, E. L., Van Beijsterveldt, C. E., & Bartels, M., Dolan, C.V., & van Bergen, E. (2019). Genetic and environmental influences on different forms of Bullying Perpetration, Bullying Victimization, and their Co-occurrence. *Behavior Genetics*. In Revision.

Abstract

Bullying comes in different forms, yet most previous genetically-sensitive studies have not distinguished between them. Given the serious consequences and the high prevalence, it is remarkable that the aetiology of bullying and its different forms has been under-researched. We present the first study to investigate the genetic architecture of bullying perpetration, bullying victimization, and their co-occurrence for verbal, physical and relational bullying. Primary-school teachers rated 8,215 twin children on bullying. For each form of bullying, we investigated, through genetic structural equation modelling, the genetic and environmental influences on being a bully, a victim or both. 34% of the children were involved as bully, victim, or both. The correlation between being a bully and being a victim varied from .59 (relational) to .85 (physical). Heritability was ~70% for perpetration and ~65% for victimization, similar in girls and boys, yet both were somewhat lower for the relational form. Shared environmental influences were modest and more pronounced among girls. The correlation between being a bully and being a victim was explained mostly by genetic factors for verbal (~71%) and especially physical (~77%) and mostly by environmental factors for relational perpetration and victimization (~60%). Genes play a large role in explaining which children are at high risk of being a victim, bully, or both. For victimization this suggests an evocative gene-environment correlation: some children are at risk of being exposed to bullying, partly due to genetically influenced traits. So, genetic influences make some children more vulnerable to become a bully, victim or both.

Introduction

Bullying in schools can take different forms: it can be direct (like name calling and hitting) and indirect (like social exclusion), but it always captures an element of power imbalance between the victim and the bully (or perpetrator). Involvement in bullying as a victim or bully, irrespective of the exact form, can have detrimental short- and long term effects (e.g., Nansel et al., 2004; Arseneault, 2018; Kretschmer, Tropf & Niezink, 2018). Both bullies and their victims suffer for example from low self-esteem, depression, and anxiety (e.g., Kaltiala-Heino, Rimpelä, Rantanen & Rimpelä, 2000; Silberg et al., 2016). There are also differences between bullies and victims. For instance, bullies tend to suffer from impulsive behavior (O'Brennan, Bradshaw & Sawyer, 2009), while victims have an increased risk of taking their own life (Gini & Espelage, 2014). In addition to children who either bully or are bullied, there are children who are both bullied and bully themselves. These bully-victims suffer most from depression and anxiety (e.g., Swearer, Song, Cary, Eagle & Mickelson, 2001). Bullying is a common phenomenon (e.g., Shetgiri, 2013), and it is important to understand why and how children differ with respect to this phenomenon. As for different forms, verbal and physical victimization are especially linked to aggression, while relational victimization (e.g., social exclusion or spreading rumors) is more associated with internalizing problems (Casper & Card, 2017), underlining the need to study them separately. The current study explores the genetic and environmental contributions to different forms of bullying perpetration (throughout this paper termed as perpetration), bullying victimization (termed as victimization), and their co-occurrence.

Being a bully or victim tends to run in families (Allison, Roeger, Smith & Isherwood, 2014; Farrington, 1993). First, with respect to victimization, Allison, Roeger, Smith and Isherwoord (2014) showed that a parents' past history of victimization is associated with an increased risk of their offspring being victimized. Whereas only 25% of the parents without a past history of being bullied reported that their offspring was victimized, in the case of parents, who had been victimized themselves, this proportion was 55%. Second, with respect to perpetration, Farrington (1993) observed a comparable inter-generational continuity. Whereas only 5.5% of the fathers who did not bully had children who bullied, 16% of the fathers who were bullies reported that their children were bullies as well. Together, these family-risk studies show that perpetration and victimization are familial, but not whether this familial transmission is genetic or environmental in nature. To determine the role of genetic and shared environmental factors, we require a genetically informative design, such as the twin design.

Several twin studies have investigated the causes of individual differences in victimization, but only one investigated the causes of individual differences in perpetration and its association with victimization. The twin studies on victimization showed mixed results. Brendgen and colleagues (2013; 2015) found a heritability of 32% in a sample of ~300 6-12 year old twin pairs using teacher-reports (2013) and a heritability of 45% in ~200 10-year-old twin pairs using self-reports (2015). Shakoor et al. (2015) reported a similar heritability estimate of 35% in a sample of ~5,000 12-year-old twin pairs using self-reports. Silberg and colleagues (2016) studied ~1,400 8-17 year old twin pairs using mother and child self-reports (combined) and reported a heritability estimate of 45%. In contrast, Ball et al. (2008) reported a higher heritability estimate of 73% in a sample of ~1,100 10-year-old twin pairs using mother-reports. Connolly and Beaver (2016) found a heritability of 70% in a sample of ~300 12-16 year old twin pairs, who reported their history of suffering repeated bullying before age 12. Bowes et al. (2013) showed that in a sample of ~1,100 twin pairs the heritability of victimization (mother and self-reports combined) was 71% in primary school and 77% in secondary school. The diverging results may be due to differences in informant (e.g., self vs. parental report), age of the participants, and (or) the type of assessment.

Moving on to perpetration, the only twin-study reported a heritability of 61% at age 10 (Ball et al., 2008). Ball et al. were also the only ones that tested whether the genetic and environmental influences on both victimization and perpetration differed in boys and girls, and found no difference. Moreover, Ball et al. (2008) looked at the co-occurrence of perpetration and victimization, which correlated .25. This correlation was found to be solely due to genetic factors common to perpetration and victimization.

The twin-studies mentioned above did not differentiate between various types of bullying, but Eastman et al. (2018) recently investigated for the first time genetic and environmental influences on different forms of victimization. The heritability estimates of self-reported verbal, physical, relational, and property victimization in early adolescence ranged from 23% for attacks on property to 42% for physical victimization. Due to limited power (N=306 pairs in the youngest of two age groups) they could not investigate whether heritability differed between boys and girls. Perpetration was not investigated.

We know that gender and the form of bullying influence prevalence rates. Specifically, most studies report that boys are more likely to be involved in bullying than girls, either as bully or victim (e.g. Nansel et al., 2001; Veldkamp et al., 2017). However, the form of bullying has a bearing on these gender-differences (e.g. Crick & Nelson, 2002; Cullerton-Sen & Crick, 2005).

Boys are more often involved in verbal (e.g. name-calling) and physical bullying (e.g. hitting), while girls are more involved in relational bullying (e.g. social exclusion). Importantly, it remains to be investigated whether genetic and environmental influences differ in boys and girls and differ between the forms of bullying. The present study is the first to investigate the genetic and environmental influences on general, verbal, physical, and relational perpetration and victimization, and on the covariance between them.

Method

Participants

Primary school teachers provided information concerning perpetration and victimization in 8,215 twins, (4,561 pairs: 1,669 MZ and 2,289 DZ). The twins were enrolled in the Netherlands Twin Register (NTR; Beijsterveldt et al., 2013), which was established by the Department of Biological Psychology at the Vrije Universiteit Amsterdam. The project was approved by the medical ethical committee of the Vrije Universiteit Amsterdam (NTR/25-05-2007). Parents of the twins, aged 7, 9, and 12 years, provided their consent to approach the teachers of the twins with a survey. Since 2010, the survey for the primary school teachers has included four items on perpetration and four items on victimization. The current study is a follow-up study of Veldkamp et al. (2017), that focused on the prevalence of perpetration and victimization, and included the same data, which were collected between 2010 and 2015. Data were excluded if (1) zygosity was unknown (*N*=193), (2) the teacher was not sufficiently familiar with the child (*N*=74), (3) the child was rated by someone other than the regular teacher (*N*=81), (4) the twins were in separate classrooms, but rated by the same teacher (*N*=11), (5) the twins were in the same classroom but rated by different teachers (*N*=108). The 8,215 twin children in the final dataset had data for at least one wave.

The data are characterized by a small degree of dependency. A subset of children had data on two (*N*=1,617) or three (*N*=93) time points, resulting in a total sample of 10,018 observations. We conducted the analyses with the complete data recognizing that the dependency may bias-down the standard errors. After removing the dependent cases and rerunning the analyses differences in the results were trivial. We also reran the analyses with the Mplus complex option, which corrects standard errors for the dependency. Again the differences in standard errors were trivial. Given lack of appreciable differences we proceeded with the original results.

Of the MZ twins, 45.1% attended separate classrooms and 54.9% the same. In the DZ twins, these figures were 49.5% and 50.5%, respectively.

Incomplete data (*N*=1,384 twin pairs) was mostly due to one of the teachers not returning the survey when the twins attended separate classrooms (*N*=1,232). The age of the children ranged from 6.52 to 12.94 years (*M*=9.48, *SD*=2.01). The degree of perpetration and victimization did hardly change with age, as indicated by correlations between age and the eight phenotypes, which ranged from -.12 to .07. Hence, age effects were not further investigated.

Measures

Perpetration and victimization

Teachers received a survey which included four items about perpetration and four matched items about victimization. Each item concerns general, physical, verbal, or relational perpetration and victimization. The 2x4 questions were scored on a five-point response scale, ranging from 0 (*never*), 1 (*once or twice*), 2 (*two or three times a month*), 3 (*about once a week*), to 4 (*several times a week*). The four items for bullying victimization were: 'How often has this student in the last couple of months… a) been bullied (in general), b) been teased, laughed at, or called names? (verbal), c) been physically bullied, such as being hit, kicked, and pushed? (physical), d) been excluded by other children, ignored, or have other students spread false rumors? (relational)'. The parts between brackets (e.g., "relational") were indeed part of the teacher items. For the original Dutch items, see the Supplementary Materials Online. Bullying perpetration was assessed with the same items, but formulated to reflect the active form (e.g. 'How often did this student in the last couple of months… a) bully other students (in general)'). Missingness at the level of the individual items was less than 1.6%.

In the case of general, verbal, and relational perpetration and victimization items, the last two response options (i.e. "*about once a week*" and "*several times a week*") were rarely chosen. Similarly, the last three response options of the physical perpetration and victimization items were rarely chosen. We therefore transformed the response scale of the general, verbal and relational items to three categories, and the response scale of the physical items to two categories.

Statistical Analyses

First, we present the prevalence of being involved in the various types of perpetration and victimization, and the phenotypic correlations. Next, we present the results of the analyses of the twin data using genetic structural equation modeling. These results include the decomposition of the phenotypic bivariate covariance matrix (perpetration – victimization) into genetic and environmental components.

Behavioral genetic analyses plan

In twin studies, we use the ACE model to decompose phenotypic variances and covariances into genetic, common and unique environmental variance components. The A (in ACE) represents additive genetic influences, the C represents environmental influences that are shared by siblings (i.e., common) and lead to similarities between them, and E represents unique environmental influences, which make siblings less alike, and measurement error. The decomposition is based on the fact that monozygotic (MZ) twins are genetically identical, while dizygotic (DZ) twins on average share 50% of the alleles that make up segregating genes. Consequently, if the MZ twin correlation is larger than the DZ twin correlation, this suggests genetic influences. If twice the DZ correlation is greater than the MZ correlation, this suggests shared environmental influences. MZ twin correlations are invariably less than one, which imply the presence of unshared or unique environmental influences and measurement error, which contribute to twin differences. In practice the decomposition is carried out by fitting the ACE model to the twin data using genetic structural equation modeling (Posthuma et al., 2003). This allows us to generalize the decomposition to multiple phenotypes. In the present case, we decompose the phenotypic 2x2 covariance matrix (perpetration – victimization by type of bullying) into 2x2 A, C and E covariance matrices. This provides information on the contributions of genetic and environmental factors to the variance of the phenotypes and to the covariance between the phenotypes. We used the bivariate Cholesky model to obtain the bivariate decomposition. This is depicted in Figure 1.

We assumed that raters may introduce systematic variation into the phenotype ratings, which reflect for example differences in raters' visions of bullying. In addition, raters who assess multiple children, can cause possible rater contrast effects. More specifically, the twins in our dataset that attend the same classroom were assessed by the same teacher. This might result in more similar bullying estimates than when the twin children were assessed by different teachers, here termed as rater effects (Bartels, Boomsma, Hudziak, van Beijsterveldt & van den Oord, 2007; Rietveld, Hudziak, Bartels, van Beijsterveldt & Boomsma, 2003). As shown in Figure 1, we included in the model a rater effect to accommodate this variation. The rater effects are assumed to contribute to the covariance between phenotype within twin members. If the twins are rated by the same teacher (i.e., twins in the same class), the rater effect may also contributed to the phenotypic covariance between twins.

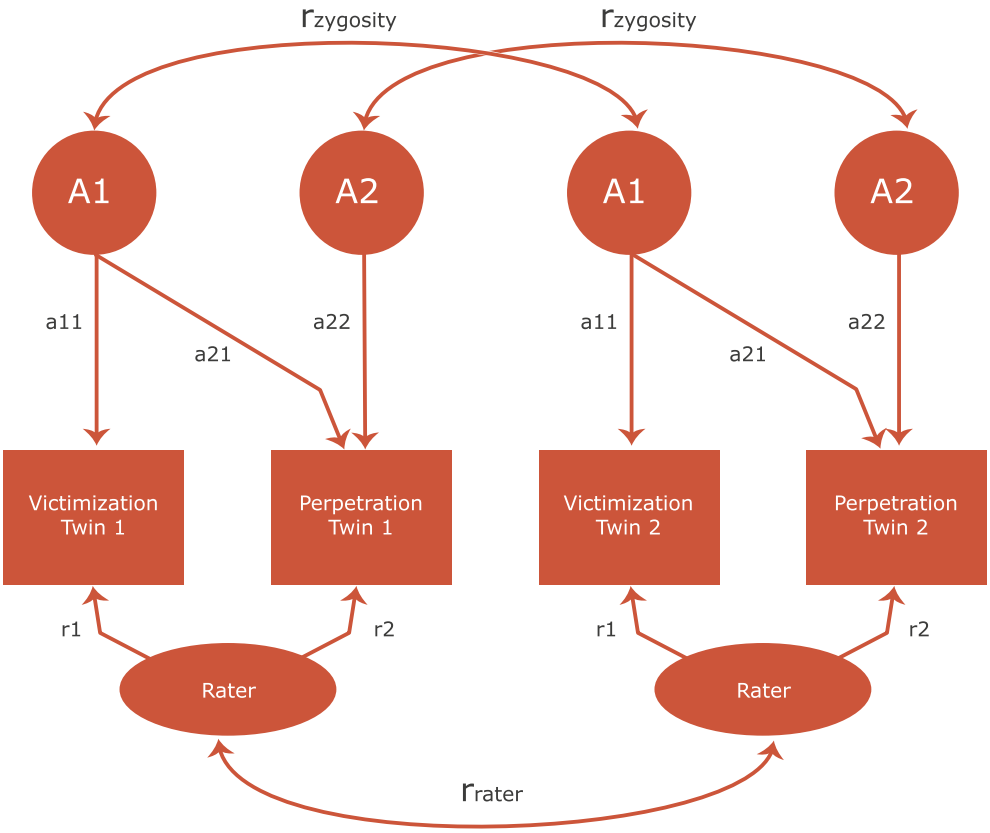


Figure 1. *Bivariate Cholesky ACE decomposition including rater bias. "A" represents the genetic influences. The common environmental (C) and unique environmental (E) influences are not shown to avoid clutter (but can be found in Figure S1 in the Supplementary Materials). "rzygosity" is 1 for MZ twins and .5 for DZ twins. "rrater" represents the correlation between the raters of the twin, which is 1 for twins rated by the same teacher and 0 for twins rated by different teachers. "a11" represents the genetic influences on victimization, "a12" represents the genetic covariance between victimization and perpetration, and "a22" represents the unique genetic influences on perpetration after accounting for the shared genetic influences. This model was fitted to each type of perpetration/victimization pair.*

We used Mplus version 7 to fit the ACE twin model (Muthén & Muthén, 1998-2012). As the data are ordinal, we used robust weighted least squared (WLS-MV) estimation applied to the tetrachoric or polychoric correlation matrices. This is consistent with the liability-threshold modeling (e.g., Rijsdijk & Sham, 2002), in which the ordinal data arise from the discretization of bivariate normal (latent) liabilities. The phenotypic summary statistics are the thresholds and the tetrachoric or polychoric correlation matrices. The correlations convey the linear association at the level of the liabilities, and the thresholds convey the frequencies of the responses. The model included 5x2 groups. First, five groups were based on zygosity and sex (MZ males, DZ males, MZ females, DZ females and DZ opposite sex). Given the five groups, we can test sex differences in the variance components and the thresholds. Second, each group was further divided into "same-class" and "different-class" groups. The latter subdivision was made to accommodate the rater effects (see Figure 1), which are shared by twins in the same class (and so the same teacher rater).

In sum, in the full model the bivariate phenotypic covariance matrix was decomposed into ACE components and the rater-variance component. The 10-group model allowed us to use likelihood ratio tests (using the DIFF test procedure in MPLUS; Muthén & Muthén, 1998-2012) to study sex differences in thresholds and variance components. We carried out the bivariate analyses (perpetration – victimization) separately for each form of bullying (general, verbal, physical and relational).

We tested the sex and classroom effects on the genetic and environmental variance components (these tests are also represented in Table S1 in the Supplementary Materials Online). First, we tested whether the thresholds (i.e., the prevalences) depended on class sharing and gender. Second, we tested whether the ACE components varied with classroom sharing and with gender. Details about model fit evaluation can be found in the Supplementary Materials.

Results

Descriptive statistics

The prevalences of children involved in perpetration and/or victimization are shown in Table 1. For these prevalence rates the response categories were dichotomized, with children scoring 0 ("never") categorized as "not involved", and children scoring 1 to 4 categorized as "involved".

Boys were more often involved in bullying, either as victim or bully. Irrespective of gender, verbal bullying was most and physical bullying was least prevalent. Regarding gender and the form of bullying, boys were more involved in verbal and physical bullying (as a bully and victim), while girls were more often involved in relational bullying (as a bully and victim).

Phenotypic correlations between all forms of perpetration and victimization are represented in Table S2, separately for boys and girls. The correlations between perpetration and victimization for the same form of bullying were for boys .64, .65, .80, and .59 for general, verbal, physical and relational, respectively, and for girls .68, .72, .85, and .68.

The model estimated twin correlations for all items are shown in Tables S3-S6. For all items the MZ correlation was higher than the DZ correlation, indicating genetic influences. The cross-twin cross-trait correlations were also all higher for MZ twins than DZ twins, suggesting that genes contribute to the perpetration-victimization association.

Table 1. *Prevalence of Victimization and Perpetration by Sex*

		Sex	
	Total Sample	**Boys**	**Girls**
Percentage victims[1]			
General	23%	27%	19%
Verbal	25%	30%	19%
Physical	8%	11%	4%
Relational	17%	15%	21%
Percentage bullies[1]			
General	26%	34%	17%
Verbal	26%	36%	17%
Physical	9%	15%	3%
Relational	20%	18%	22%

	Sex		
	Total Sample	**Boys**	**Girls**
Percentage bully-victims			
General	14%	19%	10%
Verbal	16%	21%	10%
Physical	5%	8%	2%
Relational	10%	8%	13%

Note. The percentages include children who were involved at least once or twice in the last couple of months. [1] Including bully-victims.

Bivariate genetic modeling

For each form of bullying, the same model fitting procedure was followed. Statistical details of the model fitting steps can be found in the Supplementary Materials Online.

For all forms the best fitting model was an ACE model with equal influences of genetic, common-, and unique environmental factors for twins in the same and separate classrooms. For boys and girls, the influence of genetic factors was the same, but the influence of common and unique environmental factors differed. The standardized estimates for variation due to additive genetic (A), common environmental (C) and unique environmental (E) factors, and rater estimates are shown in Table S7 and the estimates after accounting for rater effects in Table 2. Summaries of the results are visualized in Figures 2-5. All forms of bullying showed substantial genetic influences. General perpetration and general, verbal, and relational victimization showed small shared environmental influences, which were more often significant in girls. The association between perpetration and victimization was for most forms mainly genetic in nature.

Table 2. *Estimates (in %) for Variation due to Additive Genetic, Common Environmental, and Unique Environmental Factors for all Types of Perpetration, Victimization, and their Correlation, after accounting for the Rater-effects, with 95% Confidence Intervals between Brackets*

	Perpetration					Victimization					Correlation		
	A	C		E		A	C		E		*r*A	Proportion due to A[1]	
	Boys-Girls	Boys	Girls	Boys	Girls	Boys-Girls	Boys	Girls	Boys	Girls	Boys-Girls	Boys	Girls
General	72 (64-80)	9 (1-16)	13 (4-21)	19 (15-24)	16 (10-21)	62 (52-73)	9 (0-18)	13 (3-23)	29 (22-37)	25 (17-34)	.50 (.37-.62)	69 (52-87)	60 (44-77)
Verbal	73 (60-86)	2 (-9-14)	9 (-4-22)	25 (19-30)	18 (12-24)	64 (55-74)	8 (0-17)	14 (4-23)	27 (21-34)	22 (14-30)	.62 (.48-.77)	77 (63-91)	66 (53-79)
Physical	71 (54-89)	12 (-4-29)	15 (-4-34)	16 (9-24)	14 (0-28)	70 (52-87)	15 (-1-31)	18 (0-36)	15 (6-25)	13 (9-25)	.86 (.72-1.00)	81 (62-100)	73 (56-90)
Relational	68 (57-79)	7 (-4-17)	8 (-2-19)	26 (18-33)	24 (17-31)	55 (42-69)	16 (4-28)	17 (6-29)	29 (18-40)	27 (19-36)	.26 (.05-.47)	47 (11-83)	32 (6-59)

Note. Estimates for heritability were constrained to be equal for boys and girls. The rater effect was also set equal for boys and girls.

[1] The proportion of the correlation that is due to A differs between boys and girls, because the phenotypic correlations between perpetration and victimization differed.

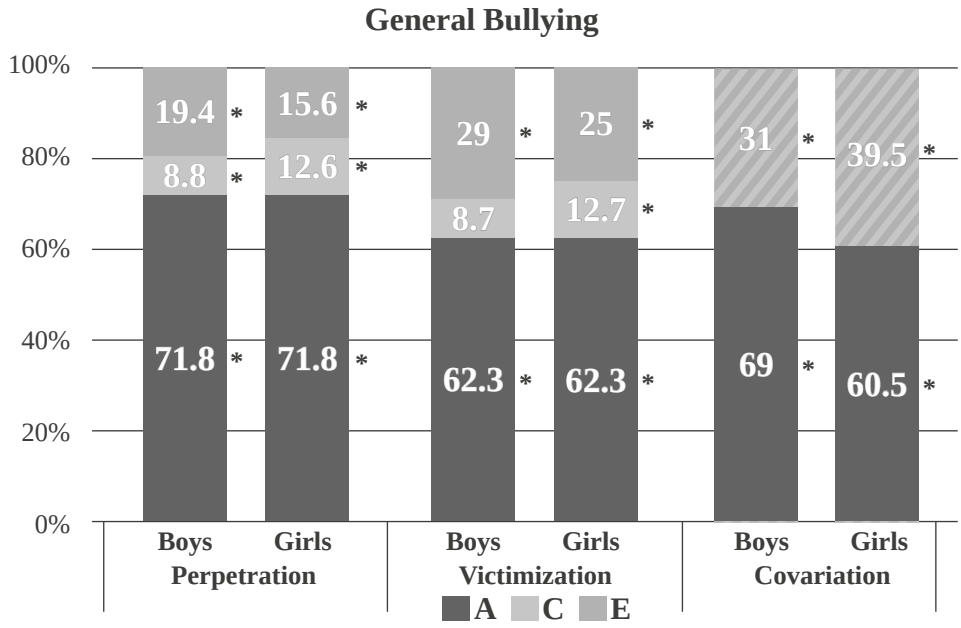


Figure 2. *Results for general bullying for Boys / Girls. The covariation is divided into shared effects (A) and environmental effects (C + E). Note that * indicates significance.*

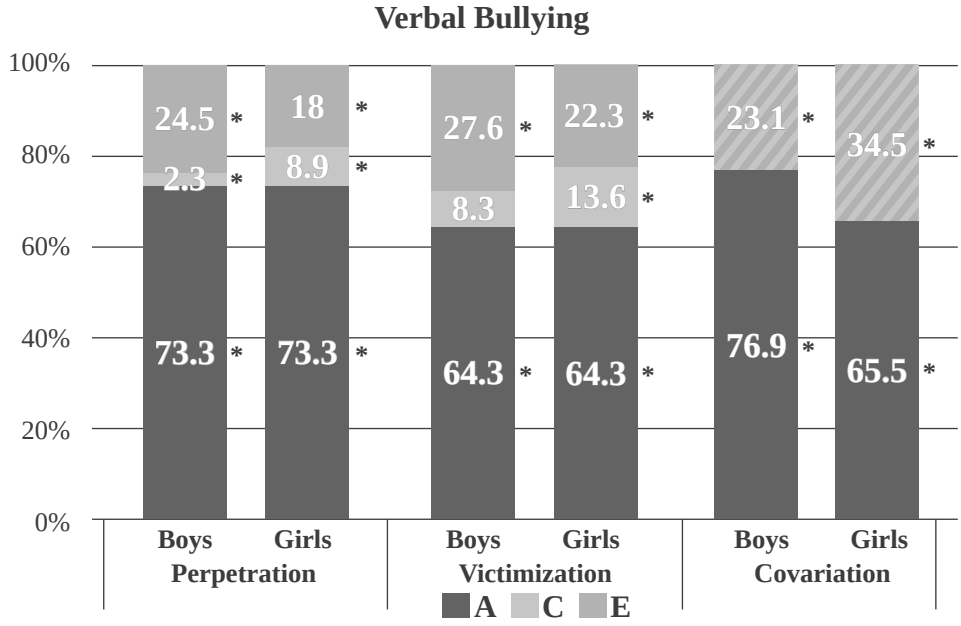


Figure 3. *Results for verbal bullying for Boys / Girls. The covariation is divided into shared effects (A) and environmental effects (C + E). Note that * indicates significance.*

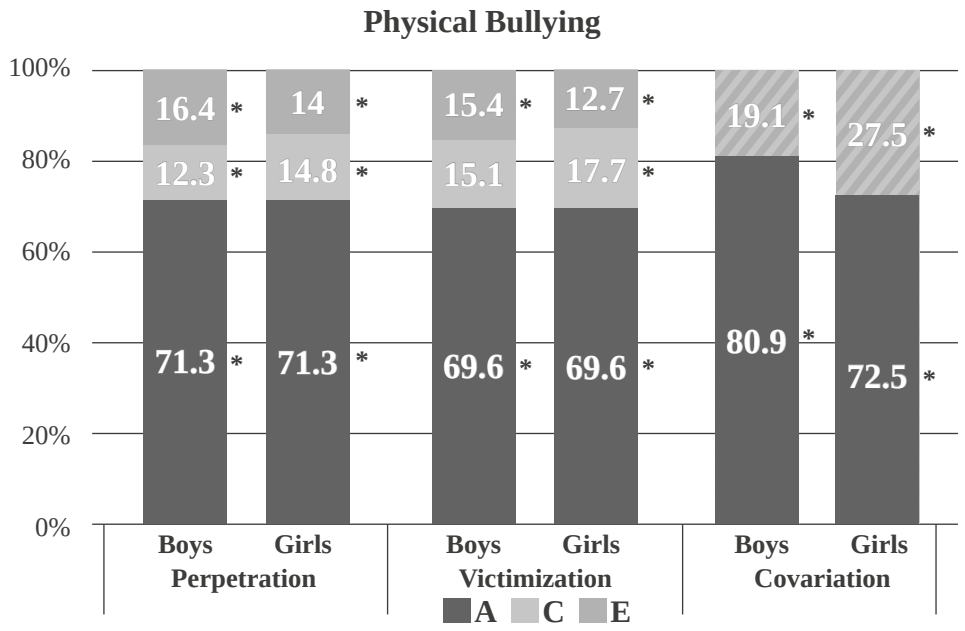


Figure 4. *Results for physical bullying for boys / girls. The covariation is divided into shared effects (A) and environmental effects (C + E). Note that * indicates significance.*

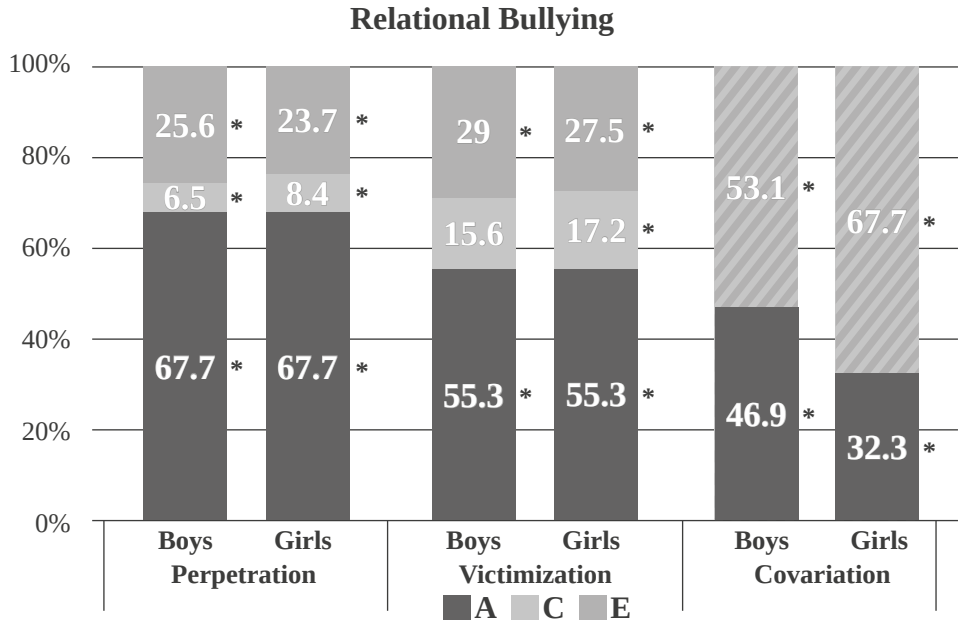


Figure 5. *Results for relational bullying for boys / girls. The covariation is divided into shared effects (A) and environmental effects (C + E). Note that * indicates significance.*

Discussion

In a sample of 8,215 primary-school children, we showed that individual differences in the liability to be a victim, bully, or bully-victim are mainly due to genetic differences between children. We asked teachers to give their view of general, verbal, physical, and relational bullying. After accounting for rater effects (twins rated by the same or different teachers), the genetic influences for both boys and girls were high for all forms of perpetration (~70%), and for general-, verbal-, and physical victimization (~65%), but somewhat lower for relational victimization (55%). The correlation between bully and victim roles was ~.70. This correlation was mostly due to shared genetic factors for the verbal and physical form and mostly due to an overlap in (common and unique) environmental factors for the relational form.

Teachers reported that the proportion of children that had been involved in bullying over the past couple of months (either as bully, victim, or both) was one third. We showed that, irrespective of gender and role (bully, victim, bully-victim), physical bullying was least prevalent and verbal bullying was most prevalent. Verbal and especially physical bullying was more common in boys, while relational bullying was more observed among girls. These prevalences provide a background for interpreting the etiological findings below.

Regarding victimization, two-thirds of the phenotypic variance expressing individual differences was due to genetic influences. At first sight, it may seem odd to claim that victimization is highly heritable, since it is an exposure to a school environment in which the child is bullied rather than direct behavior. The heritability can, however, be explained by other heritable traits that increase victimization risk. For instance, internalizing problem behavior and low self-esteem put children at greater risk to become a victim (Tsaousis, 2016) and these traits themselves are moderately heritable (Bartels et al., 2004). In addition, the risk of victimization increases with increased BMI (Janssen, Craig, Boyce & Pickett, 2004), which is highly heritable (Nan et al., 2012). Consequently, these genetically influenced traits might elicit harsh treatment by peers, leading to an evocative gene-environment correlation.

Regarding perpetration, around 70% of the individual differences were caused by genetic factors. This is slightly more than the 61% that was found in the only previous study (Ball et al., 2008). The heritability of perpetration might be easier to understand, since it is direct behavior rather than an exposure. Our finding is in line with the moderate heritability estimates of antisocial behavior (Rhee & Waldman, 2002) and aggression (Hudziak et al. al., 2003), of which bullying perpetration is one element.

It has previously been suggested that genetically influenced traits such as impulsivity could mediate the genetic effects of antisocial behavior (Jacobson, Prescott & Kendler, 2002), and this might also apply to perpetration. Bullies have indeed higher levels of impulsivity (O'Brennan, Bradshaw & Sawyer, 2009).

Being a bully or victim of physical bullying is, compared to the other subtypes, to a lesser extent affected by unique environmental factors. Unique environmental factors include factors not shared in a twin pair, as well as measurement error. Measurement error could be reduced because physical bullying is more visible for teachers than, for instance, relational bullying. Conversely, relational bullying being least heritable might be partly due to more measurement error. In accordance with this idea, Eastman et al. (2018) also showed that physical victimization is most heritable.

For all forms of bullying (both perpetration and victimization), the influence of the common environment was modest and was slightly higher for girls than for boys. About half of the common environment estimates reached statistical significance. This is in line previous mixed results: Ball et al. (2008) did not find significant influences of the common environment on perpetration and victimization, but Brendgen et al. (2008) found a significant influence on victimization. Our finding of a significant influence of the common environment on general perpetration is in line with a common environmental influence on the related phenotypes aggression and antisocial behavior (Miles & Carey, 1997). The slightly higher influence of the common environment for girls indicates that the school and/or home environment are more important for girls. To illustrate, pairs of sisters are closer than other pairs of siblings (Buist, 2010).

The co-occurrence of perpetration and victimization, reflecting bully-victims, was mainly due to genetic factors for verbal and physical bullying, but mainly due to environmental factors for relational bullying. Ball et al., (2008), the only study done so far, showed that the phenotypic correlation between perpetration and victimization was low (.25) and mostly due to genetic factors. Here we demonstrate that the influences on the co-occurrence depend on type of bullying. The genetic influences on the co-occurrence might be explained in two ways. First, it might be that the same genes influence both phenotypes via another heritable characteristic, like aggression. Bully-victims are the most aggressive group, compared to 'pure' bullies and victims (Salmivalli & Nieminen, 2002). Their genetic liability for aggression makes them more likely to get involved in a fight without necessarily a clear role as a bully or victim. Second, there might be phenotypic causality, meaning that being a bully (a genetically-mediated trait) makes a child less popular and therefore more vulnerable to become a victim as well (or vice versa). Indeed, bully-victims are the most disliked group (Veenstra et al., 2005).

In interpreting these results, it is important to mention that our results are based on teacher ratings and that phenotypes are based on only one item each. In general, teacher ratings are not highly correlated with parent and self-ratings. For perpetration, Ball et al. (2008) found a modest correlation between teacher and mother reports (r=.24). Our results may therefore present situation-specific prevalences and etiology, meaning that other influences might be responsible for school-bullying than for bullying that happens out of the sight of the teacher. For aggression, however, disagreement between teacher and mother ratings did not cause different heritability estimates (Hudziak et al., 2003). The strengths of our study include: (1) our large sample and genetically-informative design, (2) investigating subtypes of perpetration and victimization measured in the same way, (3) estimating effects free of rater effects (which was for different forms of perpetration 17-37% and for victimization 34-43%).

Some children are at risk of being exposed to bullying, partly due to genetically influenced traits, but this does not mean that bullying behavior is not modifiable. Those who work with children know that children who are outliers in some ways (e.g. behavior and appearance) are more vulnerable (Arseneault, 2018). Behavior and physical appearance are moderately to highly genetically influenced (Polderman et al., 2015). Still, bullying in schools can be reduced by creating supportive environments with evidence-based interventions (Gaffney, Ttofi & Farrington, 2018).

To conclude, this study is a first step to identify why some children are involved in different types of bullying and others are not. Our results revealed that both perpetration and victimization are substantially heritable, and that their co-occurrence is mostly due to shared genetic influences for verbal and physical bullying, but mostly due to an overlap in environmental influences for relational bullying. It must be stressed that it is certainly not someone's fate to be a bully or victim (or both), but some children are more vulnerable to these social roles, and individual differences in this vulnerability are substantially due to genetic differences. Thus, becoming a victim, bully or bully-victim is not fixed beforehand, but is not randomly determined either.

References

Allison, S., Roeger, L., Smith, B., & Isherwood, L. (2014). Family histories of school bullying: implications for parent-child psychotherapy. *AUSTRALASIAN PSYCHIATRY, 22(2), 149-153.*

Analitis, F., Velderman, M. K., Ravens-Sieberer, U., Detmar, S., Erhart, M., Herdman, M., ... & Rajmil, L. (2009). Being bullied: associated factors in children and adolescents 8 to 18 years old in 11 European countries. *PEDIATRICS, 123(2), 569-577.*

Arseneault, L. (2018). Annual Research Review: The persistent and pervasive impact of being bullied in childhood and adolescence: implications for policy and practice. *JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY, 59(4), 405-421.*

Ball, H., Arseneault, L., Taylor, A., Maughan, B., Caspi, A., & Moffitt, T. (2008). Genetic and environmental influences on victims, bullies and bully-victims in childhood. *JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY, 49(1), 104.*

Bartels, M., Boomsma, D. I., Hudziak, J. J., van Beijsterveldt, T. C., & van den Oord, E. J. (2007). Twins and the study of rater (dis) agreement. *PSYCHOLOGICAL METHODS, 12(4), 451*

Bartels, M., Boomsma, D. I., Hudziak, J. J., Rietveld, M. J., van Beijsterveldt, T. C., & van den Oord, E. J. (2004). Disentangling genetic, environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *TWIN RESEARCH AND HUMAN GENETICS, 7(2), 162-175.*

Bowes, L., Maughan, B., Ball, H., Shakoor, S., Ouellet-Morin, I., Caspi, A., ... & Arseneault, L. (2013). Chronic bullying victimization across school transitions: the role of genetic and environmental influences. *DEVELOPMENT AND PSYCHOPATHOLOGY, 25(02), 333-346.*

Brendgen, M., Girard, A., Vitaro, F., Dionne, G., Tremblay, R. E., Pérusse, D., & Boivin, M. (2013). Gene–environment processes linking peer victimization and physical health problems: a longitudinal twin study. *JOURNAL OF PEDIATRIC PSYCHOLOGY, 39(1), 96-108.*

Brendgen, M., Girard, A., Vitaro, F., Dionne, G., & Boivin, M. (2015). Gene-environment correlation linking aggression and peer victimization: do classroom behavioral norms matter? *JOURNAL OF ABNORMAL CHILD PSYCHOLOGY, 43(1), 19-31.*

Buist, K. L. (2010). Sibling relationship quality and adolescent delinquency: A latent growth curve approach. *JOURNAL OF FAMILY PSYCHOLOGY, 24(4), 400.*

Casper, D. M., & Card, N. A. (2017). Overt and relational victimization: A meta analytic review of their overlap and associations with social–psychological adjustment. *CHILD DEVELOPMENT, 88(2), 466-483.*

Connolly, E. J., & Beaver, K. M. (2016). Considering the genetic and environmental overlap between bullying victimization, delinquency, and symptoms of depression/anxiety. *JOURNAL OF INTERPERSONAL VIOLENCE, 31(7), 1230-1256.*

Craig, W., Harel-Fisch, Y., Fogel-Grinvald, H., Dostaler, S., Hetland, J., Simons-Morton, B., ... & Pickett, W. (2009). A cross-national profile of bullying and victimization among adolescents in 40 countries. *INTERNATIONAL JOURNAL OF PUBLIC HEALTH, 54(2), 216-224.*

Crick, N. R., & Nelson, D. A. (2002). Relational and physical victimization within friendships: `Nobody told me there'd be friends like these. *JOURNAL OF ABNORMAL CHILD PSYCHOLOGY, 30(6), 599-607.*

Cullerton-Sen, C., & Crick, N. R. (2005). Understanding the effects of physical and relational victimization: The utility of multiple perspectives in predicting social-emotional adjustment. *SCHOOL PSYCHOLOGY REVIEW, 34(2), 147.*

Eastman, M. L., Verhulst, B., Rappaport, L. M., Dirks, M., Sawyers, C., Pine, D. S., ... & Roberson-Nay, R. (2018). Age-Related Differences in the Structure of Genetic and Environmental *CONTRIBUTIONS TO TYPES OF PEER VICTIMIZATION. BEHAVIOR GENETICS, 48(6), 421-431.*

Farrington, D. P. (1993). Understanding and preventing bullying. In M. Tonry (Ed.), *CRIME & JUSTICE: A REVIEW OF RESEARCH (VOL. 17, PP. 381-458). CHICAGO: UNIVERSITY OF CHICAGO PRESS.*

Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2018). Evaluating the effectiveness of school-bullying prevention programs: An updated meta-analytical review. *AGGRESSION AND VIOLENT BEHAVIOR.*

Gini, G., & Espelage, D. L. (2014). Peer victimization, cyberbullying, and suicide risk in children and adolescents. *JAMA, 312(5), 545-546.*

Hudziak, J. J., van Beijsterveldt, C. E. M., Bartels, M., Rietveld, M. J., Rettew, D. C., Derks, E. M., & Boomsma, D. I. (2003). Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *BEHAVIOR GENETICS, 33(5), 575-589.*

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

Janssen, I., Craig, W. M., Boyce, W. F., & Pickett, W. (2004). Associations between overweight and obesity with bullying behaviors in school-aged children. *PEDIATRICS, 113(5), 1187-1194.*

Kaltiala-Heino, R., Rimpelä, M., Rantanen, P., Rimpelä, A. (2000). Bullying at school - an indicator of adolescents at risk for mental disorders. *JOURNAL OF ADOLESCENCE, 23, 661-674.*

Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). *NEW YORK: GUILFORD PRESS.*

Kretschmer, T., Tropf, F. C., & Niezink, N. M. (2018). Causality and pleiotropy in the association between bullying victimization in adolescence and depressive episodes in adulthood. *TWIN RESEARCH AND HUMAN GENETICS, 21(1), 33-41.*

Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture on human aggression. *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY, 72(1), 207.*

Muthén, L.K. and Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. *LOS ANGELES, CA: MUTHÉN & MUTHÉN.*

Nan, C., Guo, B., Warner, C., Fowler, T., Barrett, T., Boomsma, D., ... & Maes, H. H. (2012). Heritability of body mass index in pre-adolescence, young adulthood and late adulthood. *EUROPEAN JOURNAL OF EPIDEMIOLOGY, 27(4), 247-253.*

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *JAMA, 285(16), 2094-2100.*

Nansel, T.R., Craig, W.M., Overpeck, M.D., Saluja, G., & Ruan, W.J. (2004). Cross-national consistency in the relationship between bullying behaviours and psychosocial adjustment. *ARCHIVES OF PEDIATRICS AND ADOLESCENT MEDICINE, 158, 730–736.*

O'Brennan, L. M., Bradshaw, C. P., & Sawyer, A. L. (2009). Examining developmental differences in the social-emotional problems among frequent bullies, victims, and bully/victims. *PSYCHOLOGY IN THE SCHOOLS, 46(2), 100-115.*

Polderman, T. J., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *NATURE GENETICS, 47(7), 702.*

Posthuma D, Beem AL, de Geus EJ, van Baal GC, von Hjelmborg JB, Iachine I, Boomsma DI. Theory and practice in quantitative genetics. *TWIN RES. 2003 OCT;6(5):361-76*

R Core Team (2017). R: A language and environment for statistical computing. *R FOUNDATION FOR STATISTICAL COMPUTING, VIENNA, AUSTRIA.*

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

Rietveld, M. J., Hudziak, J. J., Bartels, M., Van Beijsterveldt, C. E. M., & Boomsma, D. I. (2003). Heritability of attention problems in children: I. cross-sectional results from a study of twins, age 3–12 years. *AMERICAN JOURNAL OF MEDICAL GENETICS PART B: NEUROPSYCHIATRIC GENETICS, 117(1), 102-113.*

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *BRIEFINGS IN BIOINFORMATICS, 3(2), 119-133.*

Salmivalli, C., & Nieminen, E. (2002). Proactive and reactive aggression among school bullies, victims, and bully victims. Aggressive Behavior: *OFFICIAL JOURNAL OF THE INTERNATIONAL SOCIETY FOR RESEARCH ON AGGRESSION, 28(1), 30-44.*

Shakoor, S., McGuire, P., Cardno, A. G., Freeman, D., Plomin, R., & Ronald, A. (2015). A shared genetic propensity underlies experiences of bullying victimization in late childhood and self-rated paranoid thinking in adolescence. *SCHIZOPHRENIA BULLETIN, 41(3), 754-763.*

Shetgiri, R. (2013). Bullying and victimization among children. *ADVANCES IN PEDIATRICS, 60(1), 33.*

Silberg, J. L., Copeland, W., Linker, J., Moore, A. A., Roberson-Nay, R., & York, T. P. (2016). Psychiatric outcomes of bullying victimization: a study of discordant monozygotic twins. *PSYCHOLOGICAL MEDICINE, 46(09), 1875-1883.*

Sulkowski, M. L., & Simmons, J. (2018). The protective role of teacher–student relationships against peer victimization and psychosocial distress. *PSYCHOLOGY IN THE SCHOOLS, 55(2), 137-150.*

Swearer, S. M., Song, S. Y., Cary, P. T., Eagle, J. W., & Mickelson, W. T. (2001). Psychosocial correlates in bullying and victimization: The relationship between depression, anxiety, and bully/victim status. *JOURNAL OF EMOTIONAL ABUSE, 2(2-3), 95-121.*

Tsaousis, I. (2016). The relationship of self-esteem to bullying perpetration and peer victimization among schoolchildren and adolescents: A meta-analytic review. *AGGRESSION AND VIOLENT BEHAVIOR, 31, 186-199.*

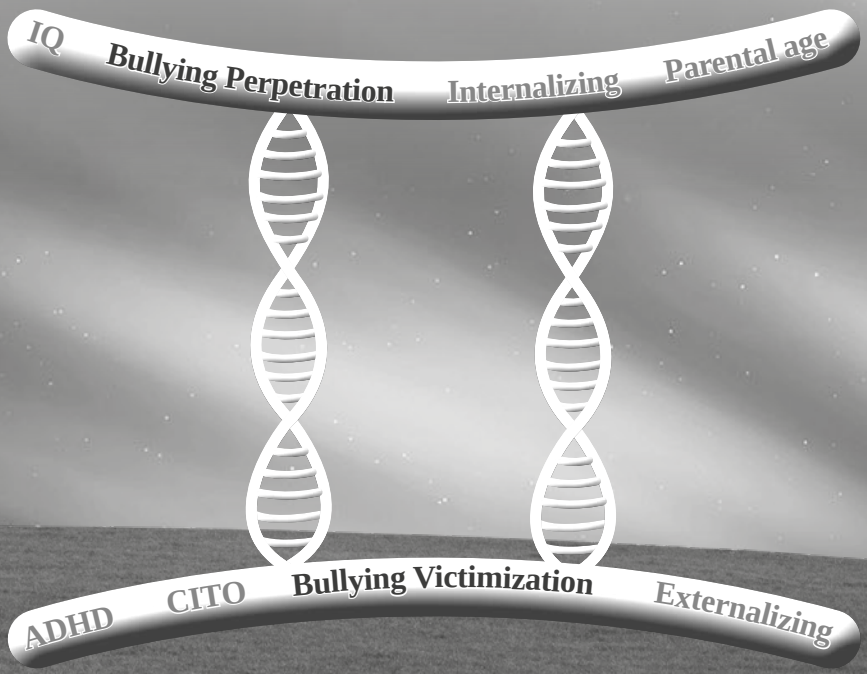
Van Beijsterveldt, C. E., Groen-Blokhuis, M., Hottenga, J. J., Franić, S., Hudziak, J. J., Lamb, D., ... & Swagerman, S. (2013). The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. *TWIN RESEARCH AND HUMAN GENETICS, 16(1), 252-267.*

Veenstra, R., Lindenberg, S., Oldehinkel, A. J., De Winter, A. F., Verhulst, F. C., & Ormel, J. (2005). Bullying and victimization in elementary schools: a comparison of bullies, victims, bully/victims, and uninvolved preadolescents. *DEVELOPMENTAL PSYCHOLOGY, 41(4), 672.*

Veldkamp, S. A.M, van Bergen, E., de Zeeuw, E. L., van Beijsterveldt, C. E., Boomsma, D.I.,& Bartels, M. (2017). Bullying and victimization: the effect of close companionship. *TWIN RESEARCH AND HUMAN GENETICS, 20(1), 19-27.*

Chapter 3 Supplement

Genetic and Environmental Influences on Different Forms of Bullying Perpetration, Bullying Victimization, and their Co-occurrence



Details about model fit evaluation

Evaluating model fit was based on the Chi-squared test, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA) (Kline, 2011). The Chi-squared test is based on the difference between the observed and expected covariance matrices. Better fit is indicated by chi-square values closer to zero. Model evaluation was based on the combinational rule of chi-squared p-values >.05, CFI >.95, and RMSEA <.05. Comparison of a model with a reduced model was based on χ^2 difference testing. To accommodate multiple testing, we used an adjusted α-value of .01. The data were prepared in R, version 3.4.1 (R Core Team, 2015) and all models were fitted in Mplus, version 7 (Muthén & Muthén, 1998-2012).

Detailed model fitting steps

For each form of perpetration and victimization, the same model fitting procedure was followed. Results of this model fitting procedure are given apart for each item.

General perpetration-victimization

The full bivariate model gave an excellent fit to the data (χ^2 (94, N = 5,634 twin pairs) = 123.66, p = .022 *CFI* = .997, *RMSEA* = .024 (90% Confidence Interval (CI): .010 - .034)). The thresholds could not be constrained to be equal for twins in same and separate classrooms ($\Delta\chi^2$(8, N = 5,634 twin pairs) = 32.825, p < .001) nor for boys and girls ($\Delta\chi^2$(8, N = 5,634 twin pairs) = 470.744, p < .001) and were therefore freely estimated in subsequent models. For twins in the same and separate classrooms, the genetic parameters ($\Delta\chi^2$(6, N = 5,634 twin pairs) = 5.25, p = .512), the common environment parameters ($\Delta\chi^2$(6, N = 5,634 twin pairs) = 11.52, p = .074) and the unique environment parameters could be constrained to be equal ($\Delta\chi^2$(2, N = 5,634 twin pairs) = 2.03, p = .363). The genetic influences did not differ significantly between for boys and girls ($\Delta\chi^2$(3, N = 5,634 twin pairs) = 5.26, p = .154), but the environmental influences (both common and unique) did ($\Delta\chi^2$(4, N = 5,634 twin families) = 21.33, p < .001). In Figure 2 of the manuscript a summary of the results is visualized.

Verbal perpetration-victimization

The full bivariate model gave a satisfactory fit to the data (χ^2(94, N = 5,610 twin pairs) = 156.94, p < .001, *CFI* = .993, *RMSEA* = .035 (90% CI: .025 - .034)). The thresholds could not be constrained to be equal for twins in same and different classrooms ($\Delta\chi^2$(8, N = 5,610 twin pairs) = 30.18, p < .001) and for boys and girls ($\Delta\chi^2$(8, N = 5,610 twin pairs) = 621.85, p < .001) and are therefore freely estimated in subsequent models.

For twins in the same and separate classrooms, the genetic parameters ($\Delta\chi^2$(6, N = 5,610 twin pairs) = 2.21, p = .899), the common environment parameters ($\Delta\chi^2$(6, N = 5,610 twin pairs) = 12.95, p = .044) and the unique environment parameters could be constrained to be equal ($\Delta\chi^2$(2, N = 5,610 twin pairs) = 2.81, p = .246). For boys and girls, the genetic can be constrained ($\Delta\chi^2$(3, N = 5,610 twin pairs) = .97, p = .810), but not the environmental influences ($\Delta\chi^2$(4, N = 5,610 twin pairs) = 21.55, p < .001). In Figure 3 of the manuscript a summary of the results is visualized.

Physical perpetration-victimization

The full bivariate model gave an excellent fit to the data (χ^2(62, N = 5,610 twin pairs) = 59.26, p = .575 *CFI* = 1.00, *RMSEA* = .000 (90% CI: .000 - .023)). The thresholds could be constrained to be equal for twins in same and separate classrooms ($\Delta\chi^2$(4, N = 5,610 twin pairs) = 2.85, p = .583), but not for boys and girls ($\Delta\chi^2$(2, N = 5,610 twin pairs) = 412.38, p < .001). Therefore, the thresholds for twin in the same and separate classrooms were set equal and only the thresholds for boys and girls were freely estimated in subsequent models. For twins in the same and separate classrooms, the genetic parameters ($\Delta\chi^2$(6, N = 5,610 twin pairs) = 2.88, p = .824), the common environmental parameters ($\Delta\chi^2$(6, N = 5,610 twin pairs) = 4.38, p = .625), and the unique environmental parameters ($\Delta\chi^2$(2, N = 5,610 twin pairs) = 7.81, p = .020) could be constrained to be equal. For boys and girls, the genetic parameters could be constrained to be equal ($\Delta\chi^2$(3, N = 5,610 twin pairs) = 3.28, p = .351), but not the environmental (both common and unique) parameters ($\Delta\chi^2$(4, N = 5,610 twin pairs) = 13.49, p = .009). In Figure 4 of the manuscript a summary of the results is visualized.

Relational perpetration-victimization

The full bivariate model gave an excellent fit to the data (χ^2(94, N = 5,611 twin pairs) = 117.59, p = .050 *CFI* = .997, *RMSEA* = .021 (90% CI: .000 - .032)). The thresholds could not be set equal for twins in same and different classrooms ($\Delta\chi^2$(8, N = 5,611 twin pairs) = 42.43, p < .001) and for boys and girls ($\Delta\chi^2$(8, N = 5,611 twin families) = 73.65, p < .001) and were therefore freely estimated in the subsequent models. The genetic influences could be constrained to be equal for twins in the same versus separate classrooms ($\Delta\chi^2$(6, N = 5,611 twin pairs) = 12.09, p = .060), as well as the common environmental effects ($\Delta\chi^2$(6, N = 5,611 twin pairs) = 4.15, p = .657) and the unique environmental effects ($\Delta\chi^2$(2, N = 5,611 twin pairs) = 7.81, p = .020). For boys and girls, the genetic factors could be constrained to be equal ($\Delta\chi^2$(3, N = 5,611 twin pairs) = 1.85, p = .605), and the environmental factors could not ($\Delta\chi^2$(4, N = 5,611 twin pairs) = 27.29, p < .001). In Figure 5 of the manuscript a summary of the results is visualized.

Dutch items Bullying Perpetration and Bullying Victimization

Hoe vaak is deze leerling in de afgelopen maanden…		niet	1 of 2 keer per week	2 of 3 keer per week	ongeveer 1 keer per week	meerdere keren per week
a.	gepest (algemeen)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b.	gepest door belediging, uitschelden, of uitlachen? (verbaal)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c.	gepest door spugen, slaan, schoppen of knijpen? (fysiek)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d.	gepest door buitensluiten, negeren of roddelen? (relationeel)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Hoe vaak heeft deze leerling in de afgelopen maanden…		niet	1 of 2 keer per week	2 of 3 keer per week	ongeveer 1 keer per week	meerdere keren per week
a.	andere leerlingen gepest (algemeen)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
b.	andere leerlingen gepest door belediging, uitschelden, of uitlachen? (verbaal)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
c.	andere leerlingen gepest door spugen, slaan, schoppen of knijpen? (fysiek)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
d.	andere leerlingen gepest door buitensluiten, negeren of roddelen?(relationeel)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Table S1. *Model-fitting Tests*

	Mean Structure
1	Classroom differences (Same Classroom vs. Different Classrooms)
2	Sex differences (Boys vs. Girls)
	Variance Components
3	Classroom differences in all genetic parameters
	(a11 a21 a22 $_{SC}$ = a11 a21 a22 $_{DC}$)
4	Classroom differences in all common environmental parameters
	(c11 c21 c22 $_{SC}$ = c11 c21 c22 $_{DC}$)
5	Classroom differences in unique environmental covariation
	(*r* e1-e2 $_{SC}$ = *r* e1-e2 $_{DC}$)
6	Sex differences in all genetic parameters
	(a11 a21 a22 $_{boys}$ = a11 a21 a22 $_{girls}$)
7	Sex differences in all common environmental parameters and in unique
	environmental covariation
	(c11 c21 c22 & *r* e1-e2 $_{boys}$ = c11 c21 c22 & *r* e1-e2 $_{girls}$)

Note. SC = Same Classroom. DC = Different Classrooms. a = genetic influences. c = common environmental influences. e = unique environmental influences.

Table S2. *Correlations between various Forms of Victimization and Perpetration by Sex*

		Victimization				Perpetration			
		General	**Verbal**	**Physical**	**Relational**	**General**	**Verbal**	**Physical**	**Relational**
Victimization	**General**	-	.85	.60	.76	**.68**	.56	.53	.47
	Verbal	.88	-	.61	.70	.59	**.72**	.59	.50
	Physical	.68	.65	-	.47	.49	.52	**.85**	.36
	Relational	.75	.71	.53	-	.54	.52	.47	**.68**
		General	**Verbal**	**Physical**	**Relational**	**General**	**Verbal**	**Physical**	**Relational**
Perpetration	**General**	**.64**	.56	.56	.47	-	.86	.68	.83
	Verbal	.54	**.65**	.56	.49	.90	-	.69	.78
	Physical	.51	.51	**.80**	.48	.73	.71	-	.51
	Relational	.35	.41	.44	**.59**	.75	.77	.56	-

Note. Correlations are shown above the diagonal for girls (grey) and below the diagonal for boys (white). Correlations between the same form of victimization and perpetration are shown in **bold** typeface.

Table S3. *Twin Correlations for General Perpetration and Victimization*

	Within Traits				Cross-Traits	
	General Perpetration		General Victimization		General Perpetration with General Victimization	
	Same Class	Different Class	Same Class	Different Class	Same Class	Different Class
MZ boys	.86 (.83-.90)	.55 (.44-.67)	.83 (.78-.87)	.42 (.27-.58)	.60 (.53-.67)	.31 (.18-.43)
MZ girls	.89 (.85-.93)	.39 (.22-.56)	.86 (.81-.92)	.34 (.17-.51)	.56 (.46-.67)	.20 (.05-.35)
DZ boys	.55 (.45-.65)	.39 (.24-.54)	.62 (.53-.71)	.24 (.07-.41)	.37 (.26-.47)	.19 (.05-.33)
DZ girls	.65 (.54-.77)	.38 (.16-.59)	.71 (.61-.80)	.14 (-.10-.39)	.55 (.43-.68)	.19 (.02-.36)
DOS	.45 (.36-.55)	.21 (.07-.35)	.52 (.43-.62)	.14 (.00-.29)	.39 (.28-.50) - .29 (.17-.41)*	.17 (.04-.30) - .09 (-.06-.25)

Note. MZ = Monozygotic twins. DZ = Dizygotic twins. DOS = Dizygotic twins of Opposite Sex.

* Correlations are between the trait in twin 1 and the trait in twin 2. The first value is the correlation between the girls' victimization score and the boys' perpetration score. The second value is the correlation between the girls' perpetration score and the boys' victimization score.

Table S4. *Twin Correlations for Verbal Perpetration and Victimization*

	Within Traits				Cross-Traits	
	Verbal Perpetration		Verbal Victimization		Verbal Perpetration with Verbal Victimization	
	Same Class	Different Class	Same Class	Different Class	Same Class	Different Class
MZ boys	.81 (.76-.86)	.57 (.45-.69)	.83 (.79-.88)	.45 (.31-.59)	.54 (.46-.63)	.32 (.19-.45)
MZ girls	.86 (.81-.91)	.56 (.38-.73)	.87 (.81-.92)	.41 (.23-.59)	.60 (.50-.70)	.47 (.32-.63)
DZ boys	.50 (.40-.60)	.26 (.11-.41)	.65 (.57-.73)	.13 (-.03-.29)	.38 (.28-.47)	.15 (.01-.28)
DZ girls	.58 (45-.70)	.27 (.03-.52)	.68 (.56-.80)	.27 (.06-.49)	.42 (.27-.56)	.20 (.03-.37)
DOS	.55 (.45-.64)	.31 (.18-.44)	.50 (.40-.60)	.16 (.02-.30)	.30 (.19-.42) - .40 (.29-.52)*	.26 (.14-.39) - .09 (-.06-.25)

Note. MZ = Monozygotic twins. DZ = Dizygotic twins. DOS = Dizygotic twins of Opposite Sex.

* Correlations are between the trait in twin 1 and the trait in twin 2. The first value is the correlation between the girls' victimization score and the boys' perpetration score. The second value is the correlation between the girls' perpetration score and the boys' victimization score.

Table S5. *Twin Correlations for Physical Perpetration and Victimization*

	Within Traits				Cross-Traits	
	Physical Perpetration		Physical Victimization		Physical Perpetration with Physical Victimization	
	Same Class	Different Class	Same Class	Different Class	Same Class	Different Class
MZ boys	.90 (.84-.96)	.51 (.31-.72)	.89 (.82-.96)	.51 (.29-.74)	.78 (.69-.87)	.51 (.34-.68)
MZ girls	.87 (.76-.99)	.69 (.35-1.00)	.93 (.85-1.00)	.60 (.23-.98)	.76 (.59-.93)	.65 (.32-.99)
DZ boys	.61 (.46-.76)	.18 (-.07-.43)	.68 (.54-.82)	.32 (.04-.60)	.51 (.36-.67)	.38 (.18-.57)
DZ girls	.66 (.39-.93)	.39 (-.14-.92)	.67 (.41-.94)	.02 (-1.00-1.00)	.63 (.41-.84)	.10 (-.74-.94)
DOS	.40 (.19-.61)	.20 (-.06-.47)	.51 (.32-.69)	.02 (-.34-.38)	.53 (.36-.70) - .22 (-.05-.48)*	.12 (-.16-.40) - .03 (-.33-.39)

Note. MZ = Monozygotic twins. DZ = Dizygotic twins. DOS = Dizygotic twins of Opposite Sex.

* Correlations are between the trait in twin 1 and the trait in twin 2. The first value is the correlation between the girls' victimization score and the boys' perpetration score. The second value is the correlation between the girls' perpetration score and the boys' victimization score.

Table S6. *Twin Correlations for Relational Perpetration and Victimization*

	Within Traits				Cross-Traits	
	Relational Perpetration		Relational Victimization		Relational Perpetration with Relational Victimization	
	Same Class	Different Class	Same Class	Different Class	Same Class	Different Class
MZ boys	.84 (.79-.89)	.48 (.31-.66)	.79 (.72-.87)	.49 (.29-.69)	.54 (.42-.66)	.17 (-.02-.35)
MZ girls	.86 (.82.91)	.32 (.14-.51)	.86 (.81-.90)	.31 (.13-.50)	.52 (.43-.62)	.28 (.12-.45)
DZ boys	.58 (.45-.70)	.31 (.13-.49)	.67 (.57-.78)	.45 (.26-.63)	.48 (.37-.60)	.04 (-.15-.22)
DZ girls	.65 (.54-.76)	.07 (-.16-.31)	.72 (.62-.82)	.08 (-.15-.32)	.54 (.42-.66)	-.05 (-.23-.14)
DOS	.51 (.40-.62)	.21 (.05-.36)	.51 (.39-.63)	.08 (-.08-.24)	.38 (.24-.51) - .37 (.24-.50)*	.09 (-.07-.24) - -.08 (-.25-.10)

Note. MZ = Monozygotic twins. DZ = Dizygotic twins. DOS = Dizygotic twins of Opposite Sex.

* Correlations are between the trait in twin 1 and the trait in twin 2. The first value is the correlation between the girls' victimization score and the boys' perpetration score. The second value is the correlation between the girls' perpetration score and the boys' victimization score.

Table S7. *Estimates (in %) for Variation due to Additive Genetic, Common Environmental, and Unique Environmental Factors for all Types of Victimization and Perpetration and their Correlation with the Rater-effects included*

	Perpetration							Victimization							Correlation		
	A		C		E	Rater		A		C		E	Rater		*r*A	Proportion due to A[1]	
	Boys-Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys-Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys-Girls	Boys	Girls
General	53	7	9	14	12	26		37	5	8	17	15	40		.50	34	32
Verbal	60	2	7	20	15	18		40	5	9	17	14	38		.62	47	42
Physical	53	9	11	12	10	26		46	10	12	10	8	34		.86	52	48
Relational	43	4	5	16	15	37		32	9	10	17	16	43		.26	16	14

Note. The genetic parameters are constrained to be equal for boys and girls, so there is only one A estimate for each item for both traits. The rater effect is equal for boys and girls as well. [1] The proportion of the correlation that is due to A differs between boys and girls, despite equal genetic parameters, because the phenotypic correlation (between perpetration and victimization) differed. The remaining part is due to environmental influences (both common and unique).

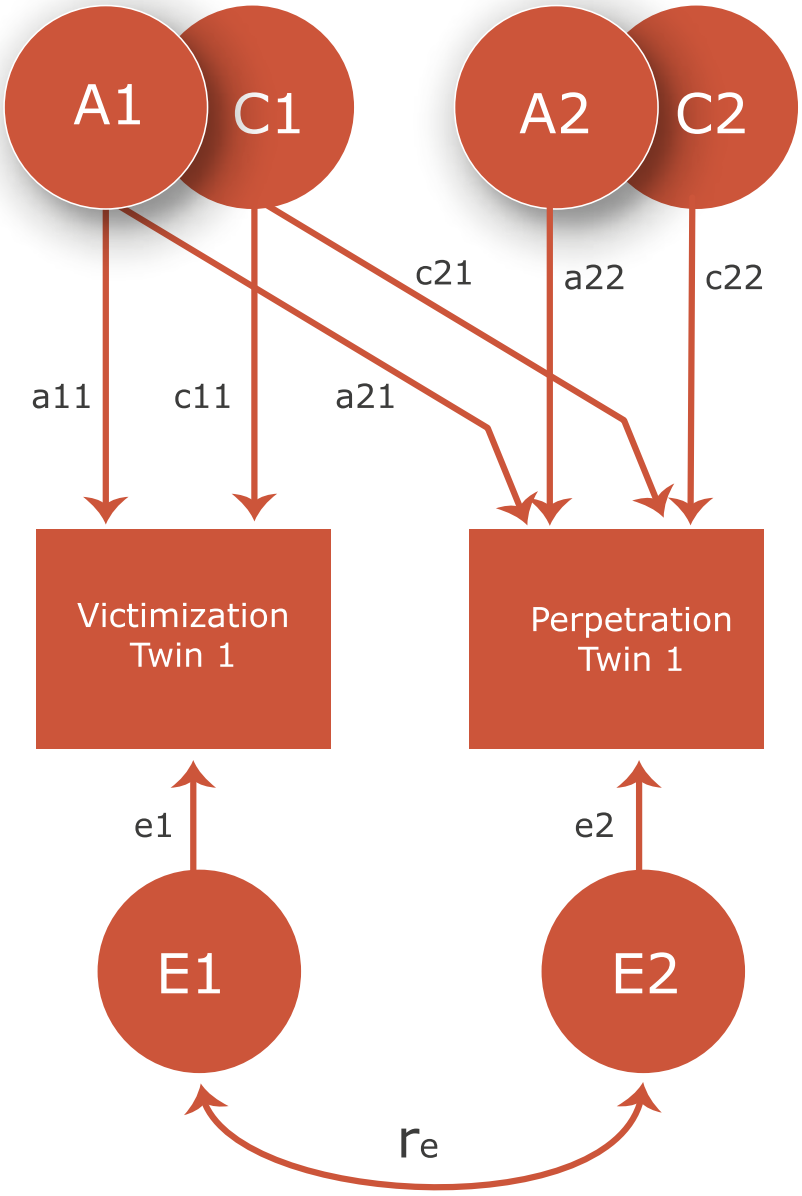

A1
C1
A2
C2
c21
a22
c22
a11
c11
a21
Victimization Twin 1
Perpetration Twin 1
e1
e2
E1
E2
re

Figure S1. *Within person specification for the ACE model used in this study. "A" represents the genetic, "C" the common environmental, and "E" the unique environmental influences. Rater effects are not shown to avoid clutter. "a11" represents the genetic influences on victimization, "a12" represents the genetic covariance between victimization and perpetration, and "a22" represents the unique genetic influences on perpetration after accounting for the shared genetic influences. "c11" represents the common environmental influences on victimization, "c12" represents the common environmental covariance between victimization and perpetration, and "c22" represents the common environmental influences on perpetration after accounting for the shared common environmental influences. "e1" represents the unique environmental influences (modeled as a residual) on victimization and "e2" the unique environmental influences (residual) on perpetration. "re" represents the unique environmental correlation between the residual of victimization and the residual of perpetration, and only this parameter for the unique environmental part of the model is estimated.*

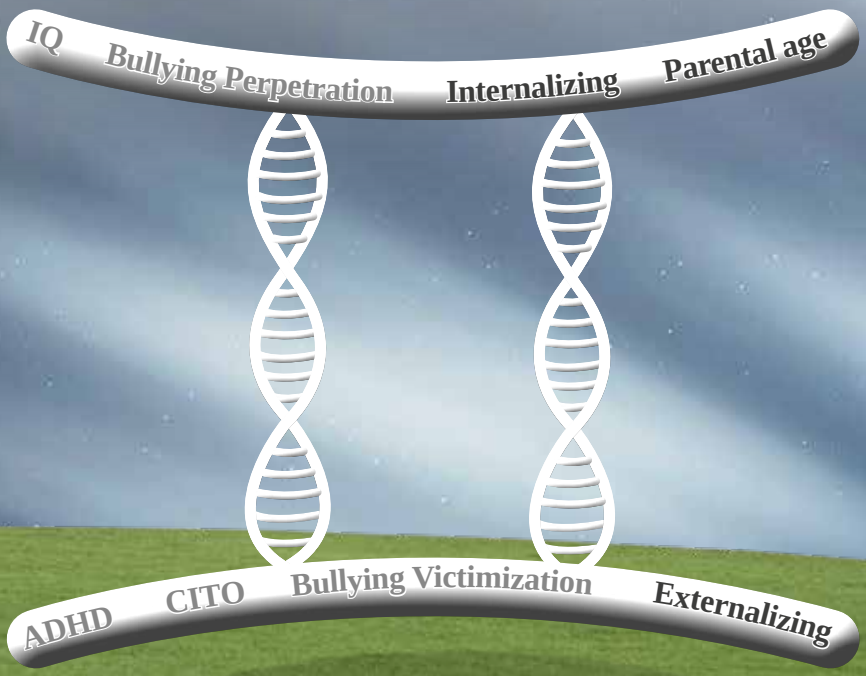
Part 2

Influences of Parental Age

Chapter 4

Parental Age and Offspring Childhood Mental Health:

A Multi-Cohort, Population-Based Investigation



This chapter is based on:

Zondervan-Zwijnenburg, M.A.J.*, Veldkamp, S.A.M. *, Neumann, A., Barzeva, S. A., Nelemans, S.A., van Beijsterveldt C.E.M., Branje, S. J. T., Hillegers, M.H.J., Meeus, W.H.J., Tiemeier, H., Hoijtink, H.J.A., Oldehinkel, A.J., & Boomsma, D.I. (2019). Parental Age and Offspring Childhood Mental Health: A Multi-Cohort, Population-Based Investigation. *Child Development*. In Press.

***These authors contributed equally to this work**

Abstract

To examine the contributions of maternal and paternal age on offspring externalizing and internalizing problems, this study analyzed problem behavior scores around age 10-12 years from four large Dutch population-based cohorts (N = 32,892). Ratings were supplied by multiple informants. Bayesian evidence synthesis was used to combine results across cohorts with 50% of the data used for discovery and 50% for confirmation. There was evidence of a robust negative linear relation between parental age and externalizing problems as reported by parents. In teacher-reports, this relation was largely explained by socio-economic status. Parental age had limited to no association with internalizing problems. Thus, in this large population-based study, no harmful effect of advanced parenthood on child problem behavior was observed.

Introduction

Since 1995, the mean maternal age at first birth has increased at a rate of 0.10 years per year in OECD countries, and in 2017 exceeded 30 years in the vast majority of these countries (Organisation for Economic Co-operation and Development, 2017). Only in Mexico was the mean age of women at childbirth lower than 28 years, and only in eight countries was it between 28 and 30 years of age. Women's reproductive years generally range from about 15 to 45 years (Te Velde, 2002). Within this wide age range some periods are generally considered more suitable to have children than others, but which parental reproductive ages are optimal for offspring physical and mental health has been a matter of debate ever since individuals have engaged in active birth control. Whereas having children at an advanced age was quite common historically, when families tended to be larger (e.g., Desjardins, Bideau, & Brunet, 1994), the current trend to delay childbearing has given rise to public health concerns.

Concerns Regarding Delayed Childbearing

Concerns regarding delayed childbearing are understandable, as a large number of research reports highlight that increased maternal age at childbirth is associated with several adverse consequences, ranging from physical problems, such as increased BMI, blood pressure and height (Carslake et al., 2017) to psychiatric conditions, such as autism (Lee & McGrath, 2015; Sandin et al., 2012), bipolar disorder (Menezes, et al., 2010), symptoms of depression, anxiety and stress (Tearne et al., 2016), and poor social functioning (Weiser et al., 2008). More recently, increased paternal age at birth has also been associated with adverse child outcomes, such as stillbirth and cleft palate (see Nybo Andersen & Urhoj, 2017, for a review). In over 40 million live births between 2007 and 2016, having an older father increased the risk of low birthweight, apgar score, and premature birth (Khandwala et al., 2018). A study of the Danish population, which included 2.8 million persons, found that older fathers are at risk of having offspring with intellectual disabilities, autism spectrum disorders and schizophrenia (McGrath et al., 2014; see De Kluiver, Buizer-Voskamp, Dolan, & Boomsma, 2017 for a review).

Several, not mutually exclusive, mechanisms have been proposed to explain the increased physical and mental health risks in offspring of older parents. First, age-related deterioration of the functioning of women's reproductive organs, such as DNA damage in germ cells, and worse quality of oocytes and placenta, can increase the risk of obstetric and perinatal complications (Myrskylä & Fenelon, 2012).

Second, male germline cells undergo cell replication cycles repeatedly during aging, with *de novo* point mutations accumulating over time (e.g., Jónsson et al., 2017) and the number of *de novo* mutations in the newborn increasing with higher age of the father at the time of conception (Kong et al. 2012; Francioli et al. 2015). Although weaker than with paternal age, *de novo* mutations in offspring correlate with maternal age as well (Goldmann et al., 2018; Wong et al., 2016). Third, genomic regions in the male germline may become less methylated with increasing age (Jenkins, Aston, Pflueger, Cairns, & Carrell, 2014) and alter the expression of health-related genes. Fourth, age effects can be due to selection, with older parents differing from younger ones in characteristics that are relevant for developmental outcomes in their offspring, such as poor social skills. The influence of selection effects can be exacerbated by assortative mating (Gratten et al., 2016). Fifth, being the child of older parents carries the risk of having to cope with parental frailty or losing a parent at a relatively young age (Myrskylä & Fenelon, 2012), and the stress evoked by these experiences may trigger health problems. Most of these mechanisms involve consequences of biological ageing. Parenthood at an advanced age is disadvantageous from a biological perspective; except for very young, physiologically immature mothers, younger parents are in a better physical condition.

Possible Benefits of Delayed Childbearing

Whereas the effects of older parental age on children's physical health and psychiatric disorders tend to be predominantly negative, the effects of older parental age on mental health problems with a stronger psychosocial component, such as externalizing and internalizing problems, tend to be more inconsistent. An indication that the negative consequences of high parental age may stretch beyond clinical diagnosis is provided by Tearne and colleagues (2015, 2016), who found that high maternal age predicted symptoms of depression, anxiety and stress in daughters, and by Janecka and colleagues (2017b) who reported a negative association between advanced paternal age and social development. In contrast, in several population-based studies, offspring of older parents, particularly of older mothers, perform better at school and work, score higher on intelligence tests, report better health and higher well-being, use fewer drugs, and have fewer behavioral and emotional problems than offspring of younger parents (e.g., Carslake, Tynelius, Van den Berg, Davey Smith, Rasmussen, 2017; McGrath et al., 2014; Myrskylä & Fenelon, 2012; Myrskylä, Barclay & Goisis, 2017; Orlebeke, Knol, Boomsma, Verhulst, 1998; Tearne et al., 2015).

While the biology of ageing seems to put older parents in an unfavorable position with regard to their offspring's physical and mental health, these contradictory effects of parental age on offspring mental health outcomes might be explained by a psychosocial perspective.

Being a child of older parents can have substantial benefits (Lawlor, Mortensen, Andersen, 2011), as older parents not only are often in a better socio-economic position than young parents (Bray, Gunnell, & Davey Smith, 2006), thereby providing a more favorable environment for children, they also have greater life experience. Furthermore, older parents display more hardiness (McMahon, Gibson, Allen, & Saunders, 2007) and tend to have less substance use and fewer mental health problems (Kiernan, 1997), hence score higher on parenting factors that promote health and development (Janecka et al., 2017a; Kiernan, 1997). In part, positive associations of advanced parental age could be related to selection effects. In young people, substance abuse and related externalizing problems go together with earlier sexual activity (Crockett, Bingham, Chopak, & Vicary, 1996), which increases the probability that intergenerational transmission of externalizing problems occurs at an early parental age (Bailey, Hill, Oesterle, & Hawkins, 2009). Like age-related parental characteristics that may have negative effects on offspring outcomes, the influence of such selection effects can be exacerbated by assortative mating (Gratten et al., 2016).

In sum, whereas advanced parenthood, particularly advanced paternal age, has primarily been associated with physical health and neurodevelopmental outcomes, such as autism and schizophrenia, advanced parenthood, particularly advanced maternal age, rather seems to predict mental health problems with a stronger psychosocial component, such as externalizing problems. Although it seems plausible that parental age interferes with subclinical problems and traits underlying these conditions, comprehensive evidence from population-based cohorts is scarce and inconsistent, and more empirical evidence is desirable. Moreover, prior population-based studies that used continuous measures of mental health problems usually focused on cognitive or behavioral problems (e.g., Carlslake et al., 2017; Orlebeke et al., 1998) and, with a few exceptions that require replication in other cohorts (Janecka et al., 2017b; Tearne et al., 2015, 2016), rarely included internalizing problems. A final reason to extend the research conducted thus far with the present study is the wide variety of populations, designs and outcomes used, which makes it hard to distinguish between substantive variation in association patterns and sample-specific artefacts. In short, there is a need for studies that investigate both maternal and paternal age effects on continuously assessed core dimensions of offspring mental health (including internalizing problems) and that use robust analytical methods which are suitable for the investigation of increased risk for both young and old parenthood.

The Present Study

We investigated parental age effects on offspring externalizing and internalizing problems around age 10-13 years in four Dutch population-based cohorts: Generation R (Gen-R), the Netherlands Twin Register (NTR), the Research on Adolescent Development and Relationships-Young cohort (RADAR-Y), and the Tracking Adolescents' Individual Lives Survey (TRAILS) (see Table 1). The Netherlands is characterized by a high maternal age at birth, and relatively few teenage pregnancies. In 1950, 1.6% of the children were born to mothers younger than 20 years of age, with a comparable percentage (1.7%) in 1990. In 2016 this number had decreased to 0.6%. In contrast, the percentages of women who gave birth at an age above 40 years were 8.5% in 1950, 1.5% in 1990, and 4.3% in 2016 (Centraal Bureau voor de Statistiek, 2018).

As the perception of childhood problems may differ for different informants (Rescorla et al., 2013; Hudziak et al., 2003), we aimed to obtain a comprehensive set of outcome measures of internalizing and externalizing problems through a multiple informant design. The four cohorts provided reports from mothers, fathers, the children themselves, and the children's teachers. The addition of reports from teachers is particularly valuable, because their reports are unlikely to be affected by parental age-related report biases. We tested both linear and nonlinear effects, to be better able to distinguish effects of older parenthood versus younger parenthood. We tested effects with and without adjusting for child gender and socio-economic status. Socio-economic status was included as a covariate to get an impression of the relative importance of socio-economic factors in explaining parental age effects.

Bayesian evidence synthesis was used to summarize the results over the cohorts. The current era is one of increased awareness of the need for replication research before making scientific claims (see, for example, Open Science Collaboration, 2015). Therefore, in this study, the datasets of the four cohort studies were used to evaluate the same set of hypotheses with respect to the relation between parental age and offspring mental health problems. This approach is called Bayesian evidence synthesis (Kuiper, Buskens, Raub, & Hoijtink, 2012).

Method

Participants

The participants in this study came from the Gen-R, NTR, RADAR-Y, and TRAILS population cohort studies. Table 2 gives the total sample size and information on parental age for each cohort.

Table 1. *General Cohort Information*

Full cohort name	Short name	Website	Birth years	References (DOI)
Generation R	Gen-R	generationr.nl	2002-2006	10.1007/s10654-016-0224-9 10.1016/j.jaac.2012.08.021
Netherlands Twin Register	NTR	tweelingenregister.org	1986-2017	10.1017/thg.2012.118 10.1016/j.jaac.2012.10.009
Research on Adolescent Development And Relationships – Young Cohort	RADAR-Y	www.uu.nl/onderzoek/radar	1990-1995	10.1111/cdev.12547 10.17026/dans-zrb-v5wp
TRacking Adolescents' Individual Lives Survey	TRAILS	trails.nl	1989-1991	10.1093/ije/dyu225

Table 2. *Cohort Descriptive Statistics of Total Sample Size and Parental Age in Current Study*

Cohort	*N*	Maternal age at child's birth child		Paternal age at child's birth child	
		Range	M (SD)	Range	M (SD)
Gen-R	4,769	16.56 – 46.85	31.68 (4.79)	17.61 – 68.67	34.24 (5.58)
NTR	25,396	17.36 – 47.09	31.35 (3.95)	18.75 – 63.61	33.76 (4.71)
RADAR-Y	497	17.80 – 48.61	31.38 (4.43)	20.34 – 52.52	33.70 (5.10)
TRAILS	2,230	16.34 – 44.88	29.32 (4.58)	18.28 – 52.09	31.99 (4.71)

The total number of children in each cohort was 4,769 for Gen-R, 25,396 for NTR, 497 for RADAR-Y, and 2,230 for TRAILS.

Gen-R mothers were recruited in the city of Rotterdam during pregnancy. Their partners, and later their children, were also invited to participate. For Gen-R, participants from the child age-10 study wave (born between 2002 and 2006) were included if they had complete information on maternal age and a child behavioral problems sum score by at least one informant. When multiple children from one family were present, one sibling was randomly removed (N = 397) to create a sample of unrelated individuals. Mean child age for mother report was: 9.72 (SD = 0.32), father report: 9.77 (SD = 0.32), and child self-report: 9.83 (SD = 0.36). 71.2% of the Gen-R sample is Dutch or European. Other ethnic groups are Suriname (6.4%), Turkish (5.3%), and Moroccan (4.2%). Mother's educational level is low (i.e., no education or primary education) for 9%, intermediate (i.e., secondary school, lower vocational training) for 42%, and high (i.e., higher vocational training, university) for 49%. Based on mother reports, 84.5% of the children had non-clinical scores for internalizing problems, 7.1% scored in the borderline category, and 8.4% scored in the clinical category. With respect to externalizing problems, 92.0% scored in the non-clinical category, 3.6% in the borderline category, and 4.5% in the clinical category.

The NTR study recruits new-born twins from all regions in the Netherlands. Here we included the data on 10-year-olds who were born between 1986 and 2008. Children were not included if they had a severe handicap which interfered with daily functioning. Mean child age for mother report was: 9.95 (SD = 0.51), father report: 9.94 (SD = 0.50) and teacher report: 9.80 (SD = 0.58). The children in NTR were mostly born in the Netherlands (99.5%). The remaining 0.5% consisted mainly of other West European nationalities (0.4%). Parents in the NTR were mostly born in the Netherlands (95.7% of fathers and 96.7% of mothers). The NTR genotype database indicates that 2.2% of participants born in the Netherlands have non-Dutch ancestry. 3.1% of mothers had a low skill occupation (primary education), 11.4% had an occupation that required lower secondary education, 40.3% had an upper secondary educational level, 30.6% had a higher vocational occupation level, and 14.6% worked at the highest (i.e. scientific) level. According to mother reports for internalizing problems, 86.1% of children had a non-clinical score, 5.9% had a borderline score, and 8.0% scored in the clinical range. For externalizing problems, 85.7% scored in the non-clinical range, 6.5% scored in the borderline range, and 7.8% in the clinical range.

The RADAR-Y sample was recruited in the province of Utrecht and four large cities in the mid–west of the Netherlands. Because the RADAR-Y study had a focus on delinquency development, children with borderline externalizing behavior problems at age 12 were oversampled. All participants from the first wave of data collection, born between 1990 and 1995, were selected. The mean age of the children at this wave was 13.03 years (*SD* = 0.46). The sample consisted mainly of native Dutch (87.9%) children. Remaining participants belonged to the following ethnic groups: Surinam (2.4%), Indonesian/ Moluccan (2.4%), Antillean (1.8%), Turkish (0.4%), and other (4.8%). Mother's educational level is low (i.e., no education or primary education) for 3.2%, intermediate (i.e., secondary school, lower vocational training) for 56.7%, and high (i.e., higher vocational training, university) for 40.1%. According to the children's reports for externalizing problems, 81.6% of the participants had a non-clinical score, 7.2% had a borderline score, and 11.2% scored in the clinical range. Using the cutoff scores for the depression scale as described by Reynolds (2004), 4.0% of the children scored in the subclinical or clinical range of depressive symptoms. Using the cutoff scores for the anxiety scale of Birmaher et al. (1997), 5.3% of the children scored in the subclinical or clinical range for anxiety symptoms.

The TRAILS sample was recruited in the Northern regions of the Netherlands. All participants from the first wave of data collection (born between 1990 and 1991) were selected. The mean age of the children at the first wave was 11.09 (*SD* = 0.56). The large majority of participants were Dutch (86.5%), with other participants being Surinam (2.1%), Indonesian (1.7%), Antillean (1.7%), Moroccan (0.7%), Turkish (0.5%), and other (6.9%). Mother's educational level is low (i.e., no education or primary education) for 6.9%, intermediate (i.e., secondary school, lower vocational training) for 66.3%, and high (i.e., higher vocational training, university) for 26.8%. Based on mother-reported sum-scores for the internalizing and externalizing scales, TRAILS participants were categorized in a non-clinical, borderline, or clinical category. For internalizing problems, 67.3% of the participants had a non-clinical score, 13.9% had a borderline score, and 18.8% had a clinical score. For externalizing problems, 74.5% had a non-clinical score, 10.2% a borderline score, and 15.4% had a score in the clinical range.

To summarize, the cohorts represented the entire Dutch geographic region across all strata from society. They had a similar distribution of SES. The percentage of participants with parents born in the Netherlands was relatively high in NTR (>95%), around 87% in Radar-Y and TRAILS, and relatively low in Gen-R (<72%). The percentage of non-clinical behavioral problem scores was lowest in TRAILS.

All studies were approved by central or institutional ethical review boards. The participants were treated in compliance with the Declaration of Helsinki, and data collection was carried out with their adequate understanding and parental consent. All measures in RADAR-Y were self-reports. In the other cohorts, children were rated by any combination of: their parents, themselves, or their teachers. Table 3 shows the total number of children in each cohort, and the number of participants with an externalizing and internalizing behavior problem score, as a function of informant (father, mother, teacher and self).

Measures

Predictors

Maternal and Paternal Age at Birth. The age of the biological parents at birth of the child was measured in years up to two decimals for each cohort.

Outcomes

Externalizing and Internalizing Problems. In most cohorts, internalizing and externalizing problems were assessed by the parent-rated Child Behavior Checklist (CBCL; Achenbach, 1991; Achenbach & Rescorla, 2001), the Youth Self-Report (YSR; Achenbach, 1991), and the Teacher Report Form (TRF; Achenbach & Rescorla, 2001). These questionnaires contain a list of around 120 behavioral and emotional problems, which can be rated as 0 = *not true*, 1 = *somewhat or sometimes true*, or 2 = *very or often true in the past 6 months*. The broadband scale Internalizing problems includes the syndromes anxious/depressed behavior, withdrawn/depressed behavior, and somatic complaints; the broadband scale Externalizing problems involves aggressive and rule-breaking behavior. In TRAILS, the Teacher Checklist of Psychopathology (TCP) was developed to be completed by teachers. The TCP contains descriptions of problem behaviors corresponding to the syndromes of the TRF. Teachers rated the TCP on a 5-point scale (De Winter et al., 2005). In Gen-R, the YSR was replaced by the Brief Problem Monitor (BPM), containing six items for internalizing and seven items for externalizing behavior problems from the YSR. All items were scored on a 3-point scale. In RADAR-Y, internalizing behavior problems were assessed by a combined score of the Reynolds Adolescent Depression Scale-2nd edition (RADS-2; Reynolds, 2000) and the Screen for Child Anxiety Related Emotional Disorders (SCARED; Birmaher, et al., 1997) questionnaires. The RADS-2 contained 23 items (the subscale anhedonia was deleted) and the SCARED contained 38 items, which were rated on a 4-point scale (1 = *almost never*, 2 = *hardly ever*, 3 = *sometimes*, 4 = *most of the time*) and 3-point scale (1 = *almost never*, 2 = *sometimes*, 3 = *often*), respectively.

Table 3. *Total Sample Size and Sample Sizes per Informant per Cohort*

(Total Sample Size) Variable	Informant	Gen-R *(N = 4,769)*		NTR *(N = 25,396)*		RADAR-Y *(N = 497)*		TRAILS *(N = 2,230)*	
Externalizing behavior problems	**Child**	BPM[a]	4,010	-	-	YSR[b]	491	YSR[b]	2,188
	Mother	CBCL[c]	4,549	CBCL[c]	21,921	-	-	CBCL[c]	1,965
	Father	CBCL[c]	3,259	CBCL[c]	14,715	-	-	-	-
	Teacher	-	-	TRF[d]	12,573	-	-	TCP[e]	1,925
Internalizing behavior problems	**Child**	BPM[a]	4,018	-	-	RADS-2[f] + SCARED[g]	266	YSR[b]	2,171
	Mother	CBCL[c]	4,550	CBCL[c]	21,731	-	-	CBCL[c]	1,955
	Father	CBCL[c]	3,259	CBCL[c]	14,626	-	-	-	-
	Teacher	-	-	TRF[d]	12,389	-	-	TCP[e]	1,924

Note. [a]Brief Problem Monitor (BPM; Achenbach, 2011).

[b]Youth Self Report (YSR; Achenbach, 1991).

[c]Child Behavior Checklist (CBCL; Achenbach, 1991; Achenbach, 2001).

[d]Teacher Report Form (TRF; Achenbach, 2001).

[e]Teacher Checklist of Psychopathology (TCP). Vignette questionnaire on the basis of the Achenbach Teacher Report Form developed by TRAILS.

[f]Reynolds Adolescent Depression Scale – 2nd edition (RADS-2; Reynolds, 2000). Excluding anhedonia scale. Standardized before averaged with SCARED.

[g]Screen for Child Anxiety Related Disorders (SCARED; Birmaher, et al., 1997). Standardized before averaged with RADS-2.

Table 3 gives an overview of the rating instruments, the informants for each of the cohorts and the number of children in each cohort for each informant/instrument combination. A sum score was calculated per informant/instrument for the relevant items for externalizing and internalizing problems respectively. Table 4 shows the mean scores for externalizing and internalizing problems per cohort. The scores for girls and boys are given in Tables S1 and S2 of the supplementary materials, respectively.

Table 4. *Mean and SD for Externalizing and Internalizing Problems*

Informant	Cohort	**Externalizing**	**Internalizing**	**N-Ext/N-Int**
Child	**Gen-R**	1.94 (1.92)	2.15 (2.09)	4,010/4,018
	RADAR-Y	10.61 (7.15)	-0.04 (0.86)	491/266
	TRAILS	8.68 (6.25)	11.28 (7.41)	2,188/2,171
Mother	**Gen-R**	3.92 (4.91)	4.86 (5.05)	4,549/4,550
	NTR	5.61 (6.12)	4.68 (5.07)	11,086/10,986
	TRAILS	8.40 (7.03)	7.85 (6.20)	1,965/1,955
Father	**Gen-R**	3.99 (4.91)	4.58 (4.72)	3,259/3,259
	NTR	4.66 (5.41)	3.56 (4.24)	7,420/7,374
Teacher	**NTR**	3.28 (5.88)	4.41 (4.96)	6,536/6,446
	TRAILS	0.44 (0.77)	0.99 (1.12)	1,925/1,924

Note. For NTR one child per family was selected to compute means and SDs. For instruments, see Note Table 3.

Covariates

Socio-Economic Status (SES) and child gender. In Gen-R, SES was defined as a continuous variable (principal component) based on parental education and household income.

In NTR, SES was a 5-level ordinal variable based on occupational level. In TRAILS, SES was a 3-level ordinal variable based on parental education, parental occupational status and household income. In RADAR-Y SES was a dichotomous variable based on parents' occupational level. Child gender was coded as male = 0 and female = 1.

Missing Data and Data Imputation

Missing Data

For externalizing problem behavior, 15.9% of the child self-reports were missing for Gen-R, while for RADAR-Y and TRAILS these percentages were 1.2% and 1.9%, respectively. For mother reported data, 4.6% were missing for Gen-R, 13.7% for NTR and 11.9% for TRAILS. For father reported data, 31.7% were missing for Gen-R and 42.1% for NTR. For teacher reported data, 50.5% were missing for NTR and 13.7% for TRAILS. For internalizing problem behavior, the percentages were similar, except for child-reported data in RADAR-Y, where 46.4% was missing. For the predictor variables, age mother and age father, 0.3% and 1.3%, were missing for NTR, 0.0% and 14.4% for Gen-R, 0.4% and 9.7% for RADAR-Y, and 5.1% and 25.0% for TRAILS, respectively. For SES, the percentage of missing values was always below 3.0%, except for Gen-R, where 22.3% was missing. For child gender, all cohorts had complete information.

Please note that the higher percentage for missing teacher- and father-reported data of NTR is due to the fact that NTR did not collect teacher-reported data at the initiation of the study and that NTR had not collected father-reported data in multiple birth years due to financial constraints. The higher percentage of missing self-reported data of internalizing problem behavior for RADAR-Y is caused by the fact that not all subscales on which the internalizing problem behavior score was based were collected from all participants.

Data Imputation

Missing data was handled by means of multiple imputation (Schafer & Graham, 2002; Van Buuren, 2012). When multiple imputation is used, the missing values are repeatedly (in this study 100 times) imputed, that is, replaced by values that are plausible given the child's scores that are not missing, resulting in 100, so-called, completed data sets. Subsequently, each completed data set is analyzed (for example, using a multiple regression) and the 100 analyses are summarized such that the fact that "artificial data" are created by imputation is properly accounted for. Multiple imputation proceeds along three steps:

1. *Determine which variables are to be used for imputation.* The variables used for imputation have to be chosen such that conditional on these variables the missing data are believed to be missing at random (MAR, van Buuren, 2012), that is, whether or not a score is missing does not depend on the missing value (Shafer & Graham, 2002). Unless missingness is planned, the variables causing the missingness are unknown to the researcher. What is often done in practice is that variables are chosen that are expected to be good predictors of the variables containing missing values. One can argue with respect to which and how many variables to use, but there is no way to test whether MAR is achieved, and MAR is an assumption. The imputation model included the outcome variables externalizing and internalizing behavioral problems per informant, total behavioral problems, SES, child gender, age of the child, age of the father and age of the mother. In some cohorts, other variables were present that could also contribute to the imputation. Specifically, parent psychopathology (in Gen-R) and total number of siblings (in NTR) contributed to the imputation model. Variables functioned only as predictors when a correlation of at least .10 with the imputed variable was present. Since the NTR dataset contained twins, the imputation process differed from that of the other cohorts. The imputation for NTR was done for each family instead of each participant, so that the same value for SES, age father and age mother was obtained for both twins. The imputation of missing data was done for informants available in each cohort. So, for example, when a cohort had no teacher-reported data, teacher data were not imputed.

2. *Generate imputed data matrices.* The R package MICE (Multiple Imputation by Chained Equations, Van Buuren & Groothuis-Oudshoorn, 2011) was used to create 100 imputed data matrices. MICE uses an iterative procedure in which sequentially each variable is imputed conditional on the real and imputed values of the other variables. Continuous variables were imputed by predictive mean matching. Categorical variables were imputed using logistic regression (see Van Buuren & Groothuis-Oudshoorn, 2011). Success of the imputation was evaluated by checking the events logged by the software, and by checking convergence plots for a lack of trends and proper mixing of the imputation chains.

3. *Analyze each imputed data set as desired and pool the results.* In the current study each of the 100 imputed data sets was analyzed using multiple regression or cluster linear regression. The results, for each regression coefficient, were 100 estimates and 100 standard errors of the estimate. As may be clear, each of the standard errors was too small because they are partly based on artificial imputed data. This was accounted for by properly pooling the results using Rubin's rules (see Van Buuren, 2012). The variance over the 100 estimates reflects the uncertainty in the estimate due to missing values (in each of the 100 completed data sets different values are imputed). In Rubin's rules the variance of the 100 estimates is used to increase the standard errors such that they properly account for the fact that part of the data is imputed. Gen-R, TRAILS and RADAR-Y used the 'pool' function of MICE in R for summarizing the effects of the 100 separate imputed datasets, whereas NTR used the pooling option of Mplus instead of R, to appropriately take into account the family clustering of the twins in the same analysis. Both pooling methods are based on the principles as explained here. The pooled estimates and standard errors were the main outcomes of the analyses after imputation.

Analytical Strategy: Bayesian Evidence Synthesis

The process of Bayesian evidence synthesis consists of four steps: (1) creating exploratory and confirmatory data sets; (2) generating competing hypotheses using exploratory analysis; (3) quantifying the support for each of the competing hypotheses using Bayesian hypothesis evaluation; and (4) Bayesian evidence synthesis, that is, summarizing the support resulting from each study into the overall support for the competing hypotheses in the data from the four cohort studies.

Exploration and Confirmation

As was elaborated in the introduction, diverse results regarding the relation between parental age and child problem behavior have been found in the literature, with increased parental age both positively and negatively related to child problem behavior. In the same vein, there may be a quadratic effect and if there is, increased child problem behavior may be present at high and low parental age. Since research is indecisive, especially for the non-clinical studies reviewed in this paper, the data resulting from each of the cohorts were split randomly into two parts containing the same number of children: an exploratory part, which was used to generate a set of competing hypotheses; and a confirmatory part, which was used to quantify the support in the data for each of the hypotheses considered. Since the NTR dataset consisted of twins, the cross-validation datasets were split based on family ID for this cohort, to ensure independent datasets. Multiple imputation was applied separately to the exploratory and confirmatory part of the data.

Having an exploratory and confirmatory dataset avoids the so-called "double dipping", that is, using the *same* data to generate and evaluate hypotheses. Here a hypothesis survived if it:1) emerged from the exploratory analyses and 2) was supported by the confirmatory analyses. The process of generating hypotheses is explained below.

Generating Hypotheses using Exploratory Analyses

The exploratory half of the data resulting from each of the four cohorts was used to generate hypotheses with respect to the relation between child problem behavior and parental age. First, for each cohort seperately, linear regression analyses were conducted to regress internalizing and externalizing problem behavior as evaluated by child, mother, father, and teacher (see Table 3 for the informants that were present per cohort) on paternal and maternal age and age squared (both with and without child gender and SES as covariates). Parental age was mean-centered to obtain the linear effect at the mean age of the samples and to reduce the correlation between the linear and quadratic terms. For Gen-R, RADAR-Y and TRAILS, the analyses were conducted in R (R Core Team, 2017). For the NTR twin-data, cluster linear regression analyses were conducted in Mplus version 8.0 (Muthén & Muthén, 1998-2017). All analyses were repeated with SES and child gender as covariates. This rendered, for each combination (e.g., predicting externalizing problems as rated by the mother from mother age and age squared) an estimate of both the linear and quadratic effect for each of the cohorts that included the informant of interest. These estimates and the corresponding p-values provided information with respect to whether the linear and non-linear effects were expected to be negative, zero, or positive. To interpret the strength of relations, the variables in the exploratory analyses were all standardized. The results of the regression analyses were translated into so-called informative hypotheses (Hoijtink, 2012), that is, hypotheses that represent expectations with respect to the state of affairs in the populations from which the data of the four cohorts were sampled. An example of such an informative hypothesis is: H_1: $\beta < 0$. That is, the regression coefficient is negative. Informative hypotheses go beyond the traditional null hypothesis (here H_0: $\beta = 0$) by stating explicitly which relations between variables are expected. Often the null is added to the set of hypotheses under consideration to protect against unjustified claims that the effect specified by an informative hypothesis exists. Another hypothesis that can be added besides the informative hypotheses is the alternative hypothesis Ha: β. That is, there are no restrictions on the regression coefficient. The alternative hypothesis is used to protect against choosing the best of a set of inadequate informative hypotheses.

For example, H_0: $\beta = 0$, and H_1: $\beta < 0$ constitute the set of hypotheses supported by the exploratory parts of the data, but both are inadequate in the confirmatory data. Instead, another unspecified hypothesis ($\beta > 0$) describes the confirmatory data best. In this case the Bayesian approach (specified below) will prefer the alternative hypothesis, H_a: β, over the informative hypotheses H_0 and H_1. By using informative hypotheses, the exact same hypotheses could be evaluated in all cohorts, even when cohorts used different measurement instruments for the same concepts. Not requiring the exact same measurement instruments is an important benefit of Bayesian evidence synthesis over classical meta-analyses.

Confirmatory Bayesian Hypotheses Evaluation

Once a set of competing informative hypotheses had been formulated (including the traditional null and alternative hypotheses), the empirical support for each pair of hypotheses was quantified using the Bayes factor (BF; Kass & Raftery, 1995). The BF is the ratio of the marginal likelihood of two competing hypotheses. Loosely speaking, the marginal likelihood of a hypothesis is the probability of that hypothesis given the data. Consequently, a BF comparing H_1 with H_a of, for example, 5 indicates that the support in the data for H_1 is five times larger than for H_a. The BF as the ratio of two marginal likelihoods implies that the fit (how well does a hypothesis describe the data set at hand) and the specificity (how specific is a hypothesis) of the hypotheses involved are accounted for (Gu, Mulder, and Hoijtink, 2018). To give an example, if $\beta = -2$, H_1: $\beta < 0$, and H_a: β, both have an excellent fit, but H_1: $\beta < 0$ is more specific than H_a: β (anything goes), and as a result, the BF will prefer H_1 over H_a. Note that the size of the BF is related to sample size. If the precision of the evidence in the data for a hypothesis increases as a result of a larger sample, the BF for that hypothesis will increase as well. The BF implemented in the R package Bain (Gu et al., 2018) was used to evaluate informative hypotheses in the context of (cluster) multiple linear regression models.

Assuming that a priori each hypothesis is equally likely to be true, the BFs were transformed in so-called posterior model probabilities (PMPs), that is, the support in the data for the hypothesis at hand given the set of hypotheses under evaluation. PMPs have values between 0 and 1 and sum to 1 for the hypotheses in the set under consideration. For example, if PMP H_0 = .05, PMP H_1 = .85, and PMP H_a = .10, then it is clear that H_1 receives the most support from the data, because it has by far the largest PMP. Thus, the result of the confirmatory Bayesian hypotheses evaluation were PMPs for each hypothesis and for each informant by each of the cohorts that had ratings by this informant. The next step was to apply Bayesian evidence synthesis.

Bayesian Evidence Synthesis

Bayesian evidence synthesis was used to summarize the support for the hypotheses of interest over the four cohort studies. Bayesian evidence synthesis (Kuiper et al., 2012) can be illustrated using the set of hypotheses: H_0: $\beta = 0$, H_1: $\beta < 0$, and H_a: β. In the context of this paper, these hypotheses are incompletely specified. The complete specification would be H_0: $\beta_1 = 0$ for NTR, H_1: $\beta_1 < 0$ for NTR and H_a: β_1 for NTR, and analogously for the other three cohort studies. This specification highlights that the support for the hypotheses depends on the cohort study at hand. Bayesian evidence synthesis can then be used to determine support for a set of hypotheses:

- H_0: H_0 for NTR & H_0 for TRAILS & H_0 for Gen-R & H_0 for Radar-Y
- H_1: H_1 for NTR & H_1 for TRAILS & H_1 for Gen-R & H_1 for Radar-Y
- H_a: H_a for NTR & H_a for TRAILS & H_a for Gen-R & H_a for Radar-Y

that is, the regression coefficient is zero *in the populations corresponding to each of the four cohort studies*, the regression coefficient is smaller than zero *in the populations corresponding to each of the four cohort studies*, and there is not prediction with respect to the regression coefficient *in the populations corresponding to each of the four cohort studies*. If for a specific set of hypotheses only two or three cohorts contain the necessary variables, the hypotheses can be adjusted accordingly. Like for each individual study, the support for these composite hypotheses was quantified using PMPs.

If a hypothesis emerges from the exploratory analyses of the data corresponding to the cohort studies and is supported by the confirmatory analyses of the data corresponding to the cohort studies, then there is evidence that this hypothesis provides an adequate description of the relation between child problem behavior and parental age, that is, in general, independent of the specific cohort studies used to evaluate this hypothesis. With the methodological approach elaborated in this section and applied in the remainder of this paper, the increased awareness of the need for replication studies before making scientific claims is explicitly addressed.

Results

Exploratory Analyses

The results of the exploratory analyses (see Supplementary Materials) generally showed a negative relation between mean-centered parental age and externalizing problems accompanied by a positive quadratic coefficient, implying that the negative relation with age at the mean declined across age (see Table S3 and Figure S1).

This model explained about 1.9% of the total variance in externalizing problems with maternal age and 1.2% with paternal age. For internalizing problems, the relation with parental age was less apparent: about 0.5% of the total variance was explained by maternal age, and about 0.2% was explained by paternal age. In analyses including the covariates SES and gender, the relation with age diminished, but remained significant (Tables S4, and S5 of the Supplementary Materials). Higher SES was related to fewer externalizing problems, and boys showed more externalizing problems than girls. In general, no relation between parental age and internalizing problems was observed (see Tables S6, S7, and S8, and Figure S1 of the Supplemental Materials).

Our interpretation of the exploratory results led to the following set of competing informative hypotheses with respect to the relation between parental age (mean-centered), as indicated by a linear (i.e., β_1) and a quadratic (i.e., β_2) coefficient, and child problem behavior:

- H_1: $\beta_1 = 0$, $\beta_2 = 0$. Age does not have a linear or quadratic relation.
- H_2: $\beta_1 < 0$, $\beta_2 = 0$. Age has a negative linear relation, there is no quadratic relation.
- H_3: $\beta_1 < 0$, $\beta_2 > 0$. Age has a negative linear relation, and a positive quadratic relation.
- H_a: β_1, β_2. The coefficients can have any value.

Based on the exploratory results, we expected most evidence for H_2 or H_3 in analyses with parental age predicting externalizing problems, and most evidence for H_1 in analyses with parental age predicting internalizing problems. Since the exploratory results did not show a positive linear or a negative quadratic relation between age and behavioral problems, the hypotheses do not include these features. However, we remained open to other options by including the alternative hypothesis Ha that imposes no constraints on the parameters, and accordingly claims that anything can be true. H_a receives the most support if none of the specified informative hypotheses provides an adequate description of the confirmatory part of the data from each of the four cohorts. In this manner, we avoided that the best hypothesis out of the set of H_1, H_2, and H_3, is an implausible hypothesis.

Confirmatory Analyses

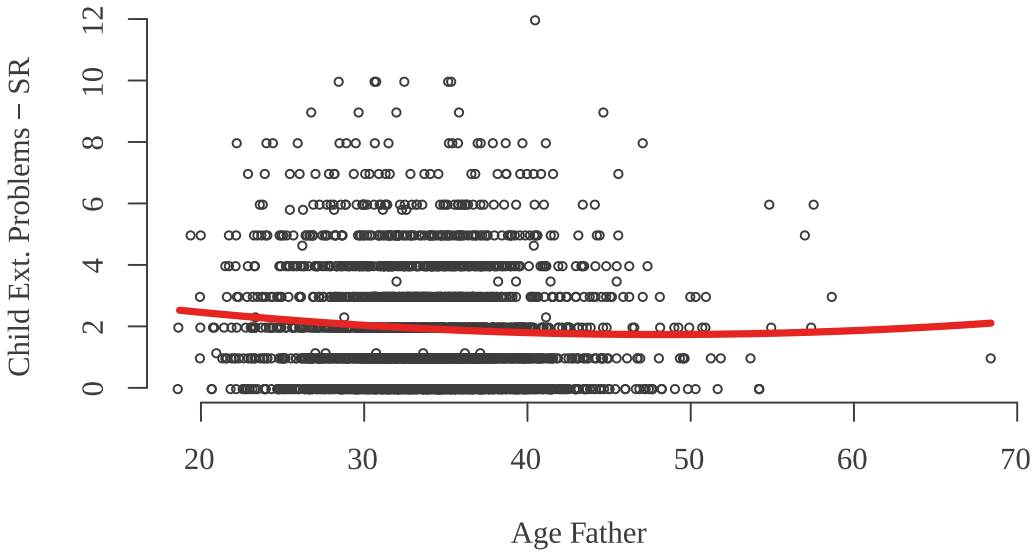
Tables S9 to S14 contain the confirmatory unstandardized regression coefficients. These are the results per cohort that generated the relative support for the competing informative hypotheses as will be presented in the next paragraph. We will discuss the underlying results briefly. Similarly to the exploratory data, the results showed negative relations across cohorts between parental age and externalizing problems. However, in the confirmatory data, the quadratic coefficients from the cohorts were less often significantly different from zero than in the exploratory data.

The model with a linear and quadratic coefficient for parental age explained on average about 1.1% of the total variance in externalizing problems with maternal age and 0.9% with paternal age as a predictor. With respect to internalizing behavior problems, the model with maternal age explained on average about 0.4% of the total variance, and paternal age explained on average about 0.3%. Figure 1 visualizes the relation between age and behavioral problems using the first imputation of the confirmatory part of Gen-R and NTR respectively. The figure presents a plot of data for internalizing and externalizing problems. As a result of centering, the linear effect that we investigated is the effect at the mean age around 29-32 years for mothers and 32-34 years for fathers (see Table 2 for mean parental age per cohort). The results presented in the figures were representative for all other analyses and cohorts.

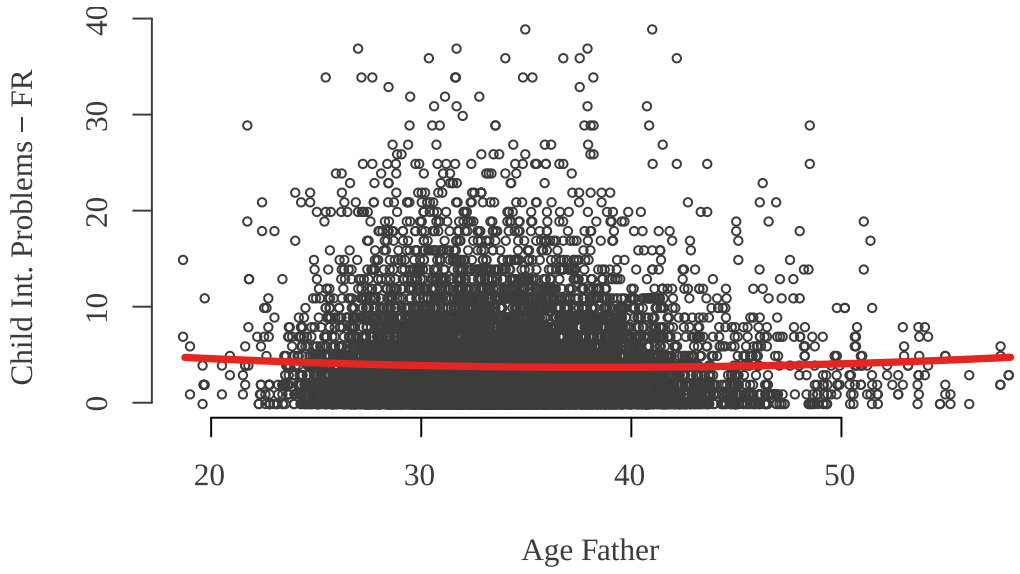
Parental Age and Externalizing Behavior Problems

The posterior model probabilities (PMPs) concerning the relation between parental age and externalizing problems are presented in Table 5. The table only shows PMP scores for those cohorts that included the associated informants (see Table 3 for an overview of informants per cohort). As shown in Table 5, for parent-reported externalizing behavior problems, Gen-R yielded most evidence for H_1 (i.e., no relation with parental age); NTR mostly supported H_2, (i.e., the relation with parental age is linear and negative) as did TRAILS, but for mother-reported externalizing behavior problems predicted by paternal age, NTR yielded most support for H_3 (i.e., the relation with parental age follows a negative linear trend including a positive quadratic factor). The combined results for mother-reported externalizing behavior problems predicted by father age showed substantial support (PMP = .53 and .45 respectively) for H_2 and H_3. For father reported externalizing behavior problems predicted by father age and for parent-reported externalizing behavior problems predicted by mother age, the combined results provided most support for H_2: the relation with parental age is linear and negative, in other words, higher parental age is associated with less externalizing behavioral problems. For teacher-reported externalizing behavior problems predicted by paternal age, TRAILS and NTR combined yielded most evidence for H_1 (i.e., no relation with parental age) closely followed by H_2. When maternal age was included, most support was found for H_2: the relation with parental age is linear and negative. For child-reported externalizing behavior problems, the results were mixed across cohorts (Gen-R preferred H_2 or H_3, RADAR-Y H_3 or H_1, and TRAILS H_1). After combining the results from the three cohorts, however, most support was obtained for H_1, that is, no relation with parental age.

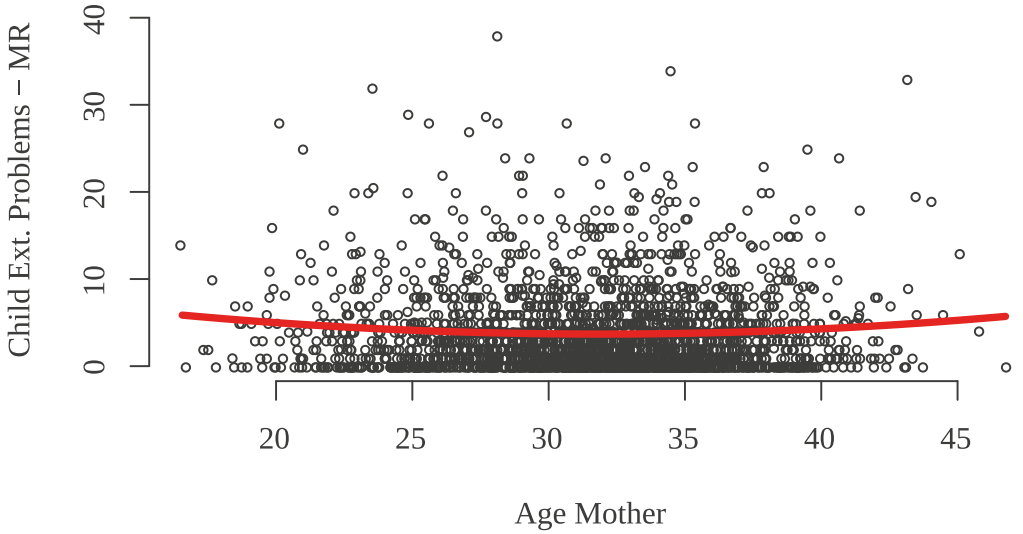
Figure 1. *Confirmatory results for parental age in relation to problem behavior as represented in Gen-R and NTR.*



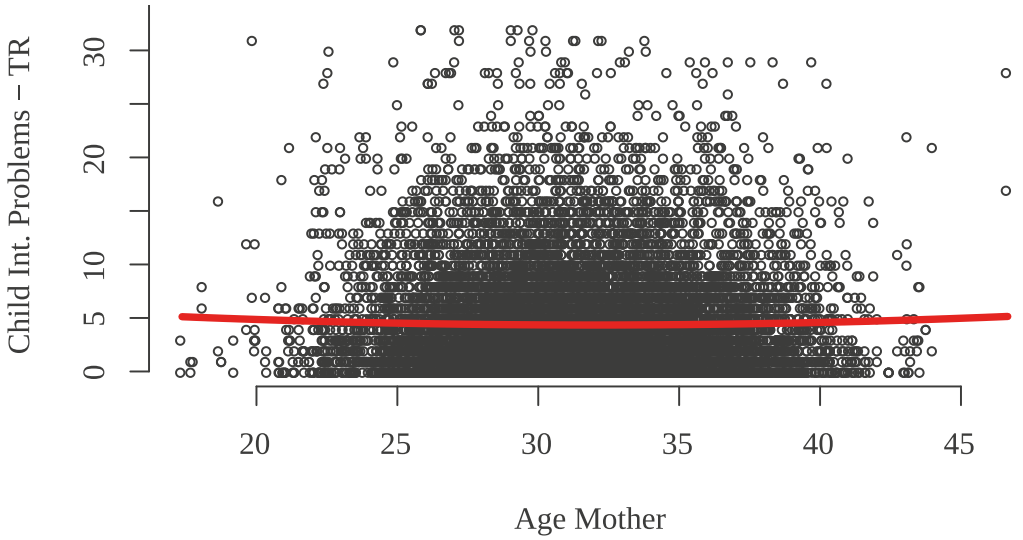
(a) Gen-R child-reported externalizing problems in relation to paternal age



(b) NTR father-reported internalizing problems in relation to paternal age



(c) Gen-R mother-reported externalizing problems in relation to maternal age



(d) NTR teacher-reported internalizing problems in relation to maternal age

Table 5. *Posterior Model Probabilities for Parental Age Predicting Externalizing Problems*

Informant	Cohort	Age Father				Age Mother			
		H_1	H_2	H_3	H_a	H_1	H_2	H_3	H_a
Child	Gen-R	.23	*.56*	.16	.05	.22	.18	*.49*	.13
	RADAR-Y	.28	.02	*.49*	.22	*.43*	.07	*.38*	.12
	TRAILS	*.86*	.13	.00	.01	*.83*	.15	.02	.01
	All	**.98**	.02	.00	.00	**.93**	.02	.04	.00
Mother	Gen-R	*.90*	.07	.02	.01	*.82*	.04	.10	.05
	NTR	.00	.02	*.74*	.24	.00	*.89*	.09	.03
	TRAILS	.18	*.74*	.06	.02	.00	*.88*	.09	.03
	All	.00	**.53**	**.45**	.00	.00	**.97**	.03	.00
Father	Gen-R	*.65*	.22	.10	.03	*.60*	.19	.17	.04
	NTR	.00	*.49*	.38	.13	.00	*.93*	.05	.02
	All	.00	**.73**	.25	.02	.00	**.95**	.05	.00
Teacher	NTR	*.55*	.41	.03	.01	.29	*.60*	.09	.02
	TRAILS	*.48*	*.31*	.16	.05	.00	*.73*	.21	.06
	All	**.67**	.32	.01	.00	.00	**.96**	.04	.00

Note. Numbers in *italic* font represent the highest posterior model probability per cohort. Numbers in **bold** font represent the highest meta-analytic results.

Table 6. *Posterior Model Probabilities for Parental Age Predicting Externalizing Problems after Correction for Impact Covariates*

Informant	Cohort	Age Father				Age Mother			
		H_1	H_2	H_3	H_a	H_1	H_2	H_3	H_a
Child	Gen-R	*.62*	.33	.04	.01	*.83*	.10	.05	.02
	RADAR-Y	*.36*	.02	*.42*	.19	*.53*	.08	.29	.10
	TRAILS	*.88*	.11	.00	.01	*.89*	.09	.02	.01
	All	**1.00**	.00	.00	.00	**1.00**	.00	.00	.00
Mother	Gen-R	*.96*	.03	.00	.00	*.97*	.02	.00	.01
	NTR	.00	.31	*.52*	.17	.00	*.95*	.04	.01
	TRAILS	*.67*	.31	.01	.01	.30	*.63*	.05	.02
	All	.03	**.99**	.00	.00	.00	**1.00**	.00	.00
Father	Gen-R	*.88*	.10	.02	.00	*.92*	.06	.01	.00
	NTR	.02	*.84*	.11	.04	.00	*.96*	.03	.01
	All	.15	**.84**	.02	.00	.00	**.99**	.01	.00
Teacher	NTR	*.79*	.20	.01	.00	*.68*	.28	.03	.01
	TRAILS	*.87*	.11	.02	.00	*.60*	.32	.07	.02
	All	**.97**	.03	.00	.00	**.81**	.18	.00	.00

Note. Numbers in *italic* font represent the highest posterior model probability per cohort. Numbers in **bold** font represent the highest meta-analytic results.

Table 6 shows the results after inclusion of the covariates as predictors of externalizing problems. After adjusting for SES and gender, all cohorts yielded substantial evidence for H_1 with respect to child- and teacher-reported externalizing problem behavior. This meant a shift especially for the child-reported problem behavior by Gen-R, and the teacher-reported problem behaviors by both NTR and TRAILS. For parent-reported problem behavior, some cohorts provided most support for H_1 (Gen-R for all parent-reports, and TRAILS for paternal age predicting mother-reported problem behavior), others for H_2 (TRAILS and NTR), and NTR for H_3 in mother-reported problem scores related to paternal age. By including covariates in the model, Gen-R and TRAILS mainly handed in support on H_2 while in NTR the support for H_2 increased at the expense of support for H_3. When combining evidence for the parent reports, most support was still found for H_2, that is, there is a linear inverse relation between parental age and externalizing problem behavior.

Parental Age and Internalizing Behavior Problems

With regard to internalizing problems (the results are presented in Table 7), the cohorts generally found most evidence for H_1 for multiple informants, except for mother-reported internalizing problems reported by maternal age in NTR. All combinations of studies rendered most support for H_1, which means that the hypothesis that there is no relation between parental age and internalizing problems was best supported by the set of studies.

After including the covariates SES and gender (Table 8), all results still suggested the most support for H_1 for the impact of parental age on internalizing problem behavior, irrespective of the cohort and informant. Consequently, combining the results from the various cohorts provided overwhelming support for H_1, that is, there is no evidence for a relation between parental age and child internalizing problem behavior.

Table 7. *Posterior Model Probabilities for Parental Age Predicting Internalizing Problems*

Informant	Cohort	Age Father				Age Mother			
		H_1	H_2	H_3	H_a	H_1	H_2	H_3	H_a
Child	**Gen-R**	*.91*	.08	.01	.00	*.86*	.09	.04	.01
	RADAR-Y	*.84*	.09	.05	.03	*.81*	.16	.02	.01
	TRAILS	*.96*	.04	.00	.00	*.93*	.06	.01	.00
	All	**1.00**	.00	.00	.00	**1.00**	.00	.00	.00
Mother	**Gen-R**	*.58*	.25	.14	.04	*.35*	.25	.33	.08
	NTR	*.69*	.26	.04	.01	.26	*.72*	.01	.01
	TRAILS	*.94*	.05	.00	.00	*.81*	.17	.02	.01
	All	**.99**	.01	.00	.00	**.71**	.29	.00	.00
Father	**Gen-R**	*.43*	*.42*	.11	.03	*.48*	.36	.13	.03
	NTR	*.96*	.04	.00	.00	*.95*	.05	.00	.00
	All	**.96**	.04	.00	.00	**.97**	.03	.00	.00
Teacher	**NTR**	*.99*	.01	.1	.00	*.99*	.01	.00	.00
	TRAILS	*.85*	.06	.07	.02	.24	.15	*.49*	.12
	All	**1.00**	.00	.00	.00	**.99**	.01	.00	.00

Note. Numbers in *italic* font represent the highest posterior model probability per cohort. Numbers in **bold** font represent the highest meta-analytic results.

Table 8. *Posterior Model Probabilities for Parental Age Predicting Internalizing Problems after Correction for Impact Covariates*

Informant	Cohort	Age Father				Age Mother			
		H_1	H_2	H_3	H_a	H_1	H_2	H_3	H_a
Child	**Gen-R**	*.77*	.21	.02	.01	*.82*	.09	.07	.02
	RADAR-Y	*.86*	.07	.04	.03	*.86*	.11	.02	.01
	TRAILS	*.97*	.03	.00	.00	*.95*	.04	.00	.00
	All	**1.00**	.00	.00	.00	**1.00**	.00	.00	.00
Mother	**Gen-R**	*.88*	.11	.01	.00	*.93*	.05	.01	.00
	NTR	*.88*	.11	.01	.00	*.70*	.29	.00	.00
	TRAILS	*.96*	.04	.00	.00	*.91*	.08	.01	.00
	All	**1.00**	.00	.00	.00	**1.00**	.00	.00	.00
Father	**Gen-R**	*.88*	.09	.02	.01	*.90*	.08	.01	.00
	NTR	*.96*	.03	.00	.00	*.96*	.04	.00	.00
	All	**1.00**	.01	.00	.00	**1.00**	.01	.00	.00
Teacher	**NTR**	*.99*	.01	.00	.00	*.99*	.01	.00	.00
	TRAILS	*.94*	.04	.02	.01	*.83*	.06	.08	.03
	All	**1.00**	.00	.00	.00	**1.00**	.00	.00	.00

Note. Numbers in *italic* font represent the highest posterior model probability per cohort. Numbers in **bold** font represent the highest meta-analytic results.

Discussion

Parental Age and Externalizing Problems

We found evidence for a negative linear relation between parental age and externalizing problems as reported by parents. That is, older parents have children with less externalizing behavior problems. There was also evidence for a negative linear relation between maternal age and externalizing problems as reported by teachers. For teachers, this finding was partly explained by SES. However, the relation between parental age and parent-reported externalizing problems persisted after adjusting for SES, so the favorable effect of parental age is not solely due to SES.

Parental Age and Internalizing Problems

Parental age seemed unrelated to child internalizing problem behavior, especially when accounting for SES. Tentatively, older parenthood might be associated with both high and low vulnerability to develop internalizing problems. On the one hand, older parents may have a lower probability of internalizing problems because they are less likely to have a background characterized by deprivation and social instability (Robson & Pevalin, 2007), known to be related to internalizing problems such as anxiety and depression. On the other hand, internalizing problems can increase the probability of older parenthood, by hampering engagement in and consolidation of romantic relationships (Manning, Trella, Lyons, & Toit, 2010; Sandberg-Thoma & Kam Dusch 2014). Possibly, both processes play a role, and their joint influence results in a lack of net result.

Sociodemographic Factors as a Potential Explanation

The relatively consistent beneficial effect of advanced parenthood for childhood externalizing problems may seem unexpected, given mixed findings from earlier research on more common mental health problems (De Kluiver, Buizer-Voskamp, Dolan, & Boomsma, 2017; McGrath et al., 2014). The beneficial effect of advanced parental age could have more than one explanation. Older and younger parents have different parenting styles. For example, there is evidence that older mothers use less frequent sanctions towards their children, are more sensitive to the child's needs and provide more structure (Trillingsgaard & Sommer, 2016). Older parents may also tend to appraise a specific problem level as less disturbing than younger parents, and older parents might be more patient and are capable of setting limits, thus feeling more equipped to handle externalizing behaviors. The positive impact of higher quality parenting by older parents is expected to be more relevant to externalizing problem behavior than to autism and schizophrenia, where a disadvantageous impact of increased parental age has been established.

Previous studies provided evidence indicating that offspring of older parents are, in several respects, more affluent than those with younger parents (e.g., Carslake et al. 2017; McGrath et al., 2014; Myrskylä & Fenelon, 2012; Orlebeke et al., 1998; Tearne et al., 2015a, 2015b). The finding that the negative relation of parental age and externalizing problems became weaker when SES was taken into account, indicates that the relatively high SES of older parents, or SES-related selection effects (Robson & Prevalin, 2007) at least partly explained why their children have a decreased probability of externalizing problems. Myrskylä, Barclay and Goisis (2017) argued that there are indeed important socio-demographic pathways associated with delayed parenthood in more recent birth cohorts. Older mothers tend to have better health behaviors during pregnancy, for example with respect to smoking during pregnancy, which is an established risk factor for offspring externalizing problems (Dolan et al., 2016).

Furthermore, parents who have externalizing behavior problems themselves may be higher in risk taking and may have children at a younger age. Hence, externalizing behavior problems may be transmitted especially by younger parents and less by older parents. This idea is in line with the unclarity about a relation between ADHD and advanced paternal age (De Kluiver et al., 2017; McGrath et al., 2014).

From a biological point of view, advanced parenthood seems mostly disadvantageous, but socio-demographic factors might compensate (or even more than compensate) for the biological disadvantages related to reproductive ageing when it comes to mental health problems. Older mothers from more recent birth cohorts are more socioeconomically advantaged, and happier after childbearing. The observation that older parents have offspring with fewer externalizing problems, tended to disappear when SES was taken into account. This shows that demographic factors can indeed compensate for the biological disadvantages.

Earlier Versus Later Birth Cohorts

In the 1950s and 1960s the number of children born to mothers over the age of 40 was larger than in 2016. For offspring born during the 1960s, Saha et al. (2009) found a negative association between maternal age and externalizing behavior problems, but in contrast to our results, they observed a positive association between maternal age and internalizing problems, and a positive association between paternal age and externalizing behavior problems. The study differed in several important aspects from the current one. All offspring were born during the 1960s, whereas in our study, all offspring were born after 1980. The age at which fathers and mothers have children has increased in the last 20 years. In the Saha et al. study average maternal and paternal ages were 24.8 and 28.4, respectively, while in our samples average maternal- and paternal ages were around 31 and 33 years.

Older mothers from earlier birth cohorts tended to have low levels of education and their offspring had many older siblings (Myrskylä, Barclay & Goisis, 2017). In later birth cohorts, older mothers had higher education than younger mothers and their offspring had fewer older siblings. Thus, the family resources are spread less thinly across siblings than in earlier times. This may be the reason that our results differ from some of the findings of Saha, Barnett, Buka and McGrath (2009). As argued by Myrskylä, Barclay & Goisis (2017), as well, being a parent during the 1960s differs from being a parent in the 1980s, and children born during the 1980s and later might benefit from positive changes in the macro-environment.

Informant Effect

We used a multi-informant design (i.e., mother, father, teacher, child) to investigate parental age effects on behavioral problems. Most questionnaires belonged to the same system (ASEBA), but they do not necessarily capture the exact same construct, as different informants observe the children in different contexts. It is well-established that correlations between different types of informants are modest at the most (Achenbach, McConaught, & Howell, 1987; Renk & Phares, 2004), and it is generally recommended to involve multiple informants to assess child and adolescent psychopathology (Jensen et al., 1999). Consistent with the notion that different informants provide partly non-overlapping information, the results in this study depended on the choice of informant, since, as opposed to parent-reported problems, child-reported externalizing problems were not predicted by parental age. Conceivably, this different outcome for child-reported problems is due to a limited ability of 10-year-old children to report reliably and validly on their externalizing behaviors. It is less likely that the associations with parent-reports are caused by reporter bias because, as teacher-reports also provided support for an association with maternal age. Thus, the choice of informant is not an arbitrary one, and may influence the associations that are found. Obviously, the parent and teacher sample sizes were also substantially larger than the sample size for child-reports. Additionally, the largest study with child reports (i.e., TRAILS) used a shortened version of the YSR, which could cause lower reliability and validity of child-reports.

Strengths of the Current Paper

This paper adopted an analysis strategy that used the data of multiple cohort studies to evaluate the same set of hypotheses. First, the data of each cohort study were divided into two parts: an exploratory part and a confirmatory part. Second, the exploratory part was used to generate a set of competing informative hypotheses. Third, the confirmatory part was used to compute the support in each cohort for the hypotheses entertained and to combine studies by means of Bayesian updating to compute overall results (Kuiper et al., 2012).

This analysis strategy had a number of advantages. In the exploratory analyses data snooping or even p-hacking is allowed, because this part of the data is only used to generate a set of competing informative hypotheses and not to evaluate these hypotheses. In contrast, the confirmatory part of each data set is only used to evaluate this set of informative hypotheses to the traditional null and alternative hypotheses, which should, especially in ages of replication crisis, publication bias and questionable research practices, increase the credibility of our results. The interested reader is referred to the Supplementary Materials where we highlight why exploratory analyses may lead to incorrect interpretations, even with large samples, and that cross-validation can prevent this from happening. In addition, with traditional null hypothesis significance testing, we would not have been able to quantify the support for the null hypothesis (*p*-values cannot be used to "accept" the null hypothesis), which appeared an important hypothesis in our study. Bayes factors and posterior model probabilities are not used to reject or not reject the null hypotheses, they are used to quantify the support in each of the cohorts for the hypotheses entertained. Furthermore, combining studies using Bayesian updating enabled us to quantify the relative evidence with respect to multiple hypotheses using the data from multiple cohorts. Again, in ages of replication crisis, it is valuable to base conclusions on data from multiple cohorts that can all be used to address the same research question.

Limitations

Although the study has a number of methodological strengths, there are also limitations. First, the study focused on children's externalizing and internalizing behavior problems and did not examine other outcomes that may be positively associated with parental age, such as physical health problems and neurodevelopmental conditions. Second, children's behavior problems were only assessed during early adolescence. Thus, the study could not investigate the possibility that the direction or magnitude of the associations may vary at different points in development. For example, previous research suggesting a negative association between parental age and individuals' well-being has focused on late adolescents and young adults (e.g., Tearne et al., 2016, Weiser et al., 2008). Third, a tiny percentage of the parents were under the age of 20 at the time of the child's birth. Although this reflects societal changes in the Netherlands, it would be important to note that some results may not replicate in other populations that have higher percentages of teenage pregnancies. This may be especially relevant when interpreting the lack of an association between parental age and children's internalizing behavior problems in this study.

Conclusion

The analytic strategy applied to large cohorts showed us a beneficial association between advanced parental age and externalizing problem behavior, while for internalizing problem behavior there was no beneficial association with parental age. We found no evidence for a harmful effect of advanced parenthood.

References

Achenbach, T. M. (1991). *MANUAL FOR THE CHILD BEHAVIOR CHECKLIST/4-18.* Burlington, VT: University of Vermont.

Achenbach, T. M. (2011). *MANUAL FOR THE ASEBA BRIEF PROBLEM MONITOR*. Burlington, VT: University of Vermont.

Achenbach, T. M., & Rescorla, L. A. (2001). *MANUAL FOR THE ASEBA SCHOOL-AGE FORMS & PROFILES.* Burlington, VT: University of Vermont.

Bailey, J., Hill, K., Oesterle, S., & Hawkins, J. (2009). Parenting practices and problem behavior across three generations: Monitoring, harsh discipline, and drug use in the intergenerational transmission of externalizing behavior. *DEVELOPMENTAL PSYCHOLOGY, 45*, 1214-1226.

Birmaher, B., Ketharpal, S., Brent, D., Cully, M., Balach, S., Kaufman, J., & Neer, S. M. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale construction and psychometric characteristics. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD AND ADOLESCENT PSYCHIATRY,* 36, 545-553. doi: 10.1097/00004583-199704000-00018.

Bray, I., Gunnell, D., & Davey Smith, G. (2006). Advanced paternal age: How old is too old? *JOURNAL OF EPIDEMIOLOGY AND COMMUNITY HEALTH,* 60, 851–853. doi: 10.1136/jech.2005.045179.

Carslake, D., Tynelius, P., Van den Berg, G., Davey Smith, G., & F., R. (2017). Associations of parental age with health and social factors in adult offspring. Methodological pitfalls and possibilities. *SCIENTIFIC REPORTS,* 7, 45278. doi: 10.1038/srep45278.

Centraal Bureau voor de Statistiek. (2018, March 30). *GEBOORTE; KERNCIJFERS.* Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37422ned/table?ts=1522410899684.

Crockett, L. J., Bingham, C. R., Chopak, J. S., & Vicary, J. R. (1996). Timing of first sexual intercourse: The role of social control, social learning, and problem behavior. *JOURNAL OF YOUTH AND ADOLESCENCE*, 25, 89-111. doi:10.1007/BF01537382.

De Kluiver, H., Buizer-Voskamp, J. E., Dolan, C. V., & Boomsma, D. I. (2017). Paternal age and psychiatric disorders: A review. *AMERICAN JOURNAL OF MEDICAL GENETICS PART B: NEUROPSYCHIATRIC GENETICS*, 174, 202–213. doi: 10.1002/ajmg.b.32508.

De Winter, A. F., Oldehinkel, A. J., Veenstra, R., Brunnekreef, J. A., Verhulst, F. C., & Ormel, J. (2005). Evaluation of nonresponse bias in mental health determinants and outcomes in a large sample of preadolescents. *EUROPEAN JOURNAL OF EPIDEMIOLOGY*, 20, 173-181.

Desjardins, N., Bideau, A., & Brunet, G. (1994). Age of mother at last birth in two historical populations. *JOURNAL OF BIOSOCIAL SCIENCE*, 26, 509-516. doi: 10.1017/S0021932000021635

Dolan, C. V., Geels, L., Vink, J. M., van Beijsterveldt, C. E. M., Neale, M. C., Bartels, M., & Boomsma, D. I. (2016). Testing causal effects of maternal smoking during pregnancy on offspring's externalizing and internalizing behavior. *BEHAVIOR GENETICS*, 46, 378-388. doi: 10.1007/s10519-015-9738-2

Francioli, L. C., Polak, P. P., Koren, A., Menelaou, A., Chun, S., Renkens, I., ... & Slagboom, P. E. (2015). Genome-wide patterns and properties of de novo mutations in humans. *NATURE GENETICS*, 47, 822. doi: 10.1038/ng.3292

Goisis, A., Schneider, D. C., & Myrskylä, M. (2017). The reversing association between advanced maternal age and child cognitive ability: evidence from three UK birth cohorts. *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY*, 46, 850-859. doi: 10.1093/ije/dyw354

Goisis, A., Schneider, D. C., & Myrskylä, M. (2018). Secular changes in the association between advanced maternal age and the risk of low birth weight: a cross-cohort comparison in the UK. *POPULATION STUDIES*, 72, 381-397. doi: 10.1080/00324728.2018.1442584

Goldmann, J. M., Seplyarskiy, V. B., Wong, W. S., Vilboux, T., Neerincx, P. B., Bodian, D. L., ... & Niederhuber, J. E. (2018). Germline de novo mutation clusters arise during oocyte aging in genomic regions with high double-strand-break incidence. *NATURE GENETICS*, 50, 487-492. doi: 10.1038/s41588-018-0071-6

Gratten J., Wray N.R., Peyrot, W.J., McGrath, J.J., Visscher, P.M., Goddard, M.E. (2016). Risk of psychiatric illness from advanced paternal age is not predominantly from de novo mutations. *NATURE GENETICS*, 48, 718-24. doi: 10.1038/ng.3577

Gu, X., Mulder, J., & Hoijtink, H. (2018). Approximated adjusted fractional Bayes factors: A general method for testing informative hypotheses. *BRITISH JOURNAL OF MATHEMATICAL AND STATISTICAL PSYCHOLOGY*, 71, 229-261. doi: 10.1111/bmsp.12110.

Hoijtink, H. (2012). I*NFORMATIVE HYPOTHESES: THEORY AND PRACTICE FOR BEHAVIORAL AND SOCIAL SCIENTISTS.* Boca Raton, FL: CRC Press.

Hudziak, J. J., Van Beijsterveldt, C. E. M., Bartels, M., Rietveld, M. J., Rettew, D. C., Derks, E. M., & Boomsma, D. I. (2003). Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *BEHAVIOR GENETICS*, 33, 575-589. doi: 10.1023/A:1025782918793

Janecka, M., Haworth, C. M., Ronald, A., Krapohl, E., Happé, F., Mill, J., . . . Rijsdijk, F. (2017b). Paternal age alters social development in offspring. J*OURNAL OF THE AMERICAN ACADEMY OF CHILD AND ADOLESCENT PSYCHIATRY*, 56, 383-390. doi: 10.1016/j.jaac.2017.02.006

Janecka, M., Rijsdijk, F., Rai, D., Modabbernia, A., & Reichenberg, A. (2017a). Advantageous developmental outcomes of advancing paternal age. *TRANSLATIONAL PSYCHIATRY*, 7, e1156. doi:10.1038/tp.2017.125.

Jenkins, T. G., Aston, K. I., Pflueger, C., Cairns, B. R., & Carrell, D. T. (2014). Age-associated sperm DNA methylation alterations: Possible implications in offspring disease susceptibility. *PLOS GENETICS, 10*, e1004458. doi: 10.1371/journal.pgen.1004458

Jónsson, H., Sulem, P., Kehr, B. K., Zink, F., Hjartarson, E., Hardarson, M. T., . . . Stefansson, K. (2017). Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *NATURE*, 549, 519-522. doi: 10.1038/nature24018

Kass, R., & Raftery, A. (1995). Bayes Factors. *JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION*, 90, 773-795. doi: 10.1080/01621459.1995.10476572.

Khandwala, Y. S. Baker, V. L., Shaw, G. M. Stevenson, D. K., Lu, Y., & Eisenberg, M. L. (2018). Association of paternal age with perinatal outcomes between 2007 and 2016 in the United States: Population based cohort study. *BMJ, 363(K4372)*. doi: 10.1136/bmj.k4372

Kiernan, K. E. (1997). Becoming a young parent: A longitudinal study of associated factors. *BRITISH JOURNAL OF SOCIOLOGY,* 48, 406-428. doi: 10.2307/591138

Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., ... & Wong, W. S. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *NATURE*, 488, 471-475. doi: 10.1038/nature11396

Kuiper, R., Buskens, V., Raub, W., & Hoijtink, H. (2012). Combining statistical evidence from several studies: A method using Bayesian updating and an example from research on trust problems in social and economic exchange. *SOCIOLOGICAL METHODS & RESEARCH*, 42, 60-81. doi: 10.1177/0049124112464867.

Lawlor, D. A., Mortensen, L., & Andersen, A. M. (2011). Mechanisms underlying the associations of maternal age with adverse perinatal outcomes: A sibling study of 264 695 Danish women and their firstborn offspring. *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY,* 40, 1205-1214. doi: 10.1093/ije/dyr084

Lee, B. K., & McGrath, J. J. (2015). Advancing parental age and autism: Multifactorial pathways. *TRENDS IN MOLECULAR MEDICINE*, 21, 118-125. doi: 10.1016/j.molmed.2014.11.005

Manning, W. D., Trella, D., Lyons, H., & Toit, N. C. (2010). Marriageable women: A focus on participants in a community healthy marriage program. *FAMILY RELATIONS*, 59, 87-102. doi: 10.1111/j.1741-3729.2009.00588.x

McGrath, J. J., Petersen, L., Agerbo, E., Mors, O., Mortensen, P. B., & Pedersen, C. B. (2014). A comprehensive assessment of parental age and psychiatric disorders. *JAMA PSYCHIATRY,* 71, 301-309. doi: 10.1001/jamapsychiatry.2013.4081

McMahon, C., Gibson, F. L., Allen, J. L., & Saunders, D. (2007). Psychosocial adjustment during pregnancy for older couples conceiving through assisted reproductive technology. *HUMAN REPRODUCTION,* 22, 1168-1174. doi: 10.1093/humrep/del502

Menezes, P. R., Lewis, G., Rasmussen, F., Zammit, S., Sipos, A., Harrison, G., . . . Gunnell, D. (2010). Paternal and maternal ages at conception and risk of bipolar affective disorder in their offspring. *PSYCHOLOGICAL MEDICINE*, 40, 477-485. doi: 10.1017/S003329170999064X

Muthén, L., & Muthén, B. O. (1998-2017). *MPLUS USERS'* guide (8 ed.). Los Angeles, CA.

Myrskylä, M., Barclay, K., & Goisis, A. (2017). Advantages of later motherhood. Der Gynäkologe, 50, 767-772. doi: 10.1007/s00129-017-4124-1

Myrskylä, M., & Fenelon, A. (2012). Maternal age and offspring adult health: Evidence from the Health and Retirement Study. *DEMOGRAPHY, 49*, 1231–1257. doi: 10.1007/s13524-012-0132-x.

Nybo Andersen, A. M., & Urhoj, S. (2017). Is advanced paternal age a health risk for the offspring? *FERTILITY AND STERILITY,* 107, 312–318. doi: 10.1016/j.fertnstert.2016.12.019.

OECD. (2017). OECD *FAMILY DATABASE.* Retrieved Oct 6, 2017, from Family Database: http://www.oecd.org/els/family/database.htm

Oldehinkel, A. J., Rosmalen, J. G., Buitelaar, J. K., Hoek, H. W., Ormel, J., Raven, D., . . . Hartman, C. A. (2015). Cohort profile update: The TRacking Adolescents' Individual Lives Survey (TRAILS). *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY*, 44, 76-76n. doi: 10.1093/ije/dyu225.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *SCIENCE,* 349, 6251. doi :10.1126/science.aac4716

Orlebeke, J. F., Knol, D. L., Boomsma, D. I., & Verhulst, F. C. (1998). Frequency of parental report of problem behavior in children decreases with increasing maternal age at delivery. *PSYCHOLOGICAL REPORTS*, 82, 395-404. doi: 10.2466/pr0.1998.82.2.395

R Core Team. (2017). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Rescorla, L. A., Ginzburg, S., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., ... & Döpfner, M. (2013). Cross-informant agreement between parent-reported and adolescent self-reported problems in 25 societies. *JOURNAL OF CLINICAL CHILD & ADOLESCENT PSYCHOLOGY*, 42, 262-273. doi: 10.1080/15374416.2012.717870

Robson, K., & Pevalin, D. J. (2007). Gender differences in the predictors and socio-economic outcomes of young parenthood in Great Britain. *RESEARCH IN SOCIAL STRATIFICATION AND MOBILITY,* 25, 205-218. doi: 10.1016/j.rssm.2007.08.002

Reynolds, W. M. (2000). Reynolds Adolescent Depression Scale – 2nd edition (RADS-2). *PROFESSIONAL MANUAL*. Lutz, FL: Psychological Assessment Resources.

Saha, S., Barnett, A. G., Buka, S. L., & McGrath, J. J. (2009). Maternal age and paternal age are associated with distinct childhood behavioural outcomes in a general population birth cohort. *SCHIZOPHRENIA RESEARCH*, 115, 130-135. doi: 10.1016/j.schres.2009.09.012

Sandberg-Thoma, S. E., & Kam Dusch, C. M. (2014). Indicators of adolescent depression and relationship progression in emerging adulthood. *JOURNALOF MARRIAGE AND FAMILY*, 76, 191–206. doi: 10.1111/jomf.12081

Sandin, S., Hultman, C. M., Kolevzon, A., Gross, R., MacCabe, J. H., & Reichenberg, A. (2012). Advancing maternal age is associated with increasing risk for autism: A review and meta-analysis. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY*, 51, 477-486. doi: 10.1016/j.jaac.2012.02.018

Schafer, J.L. & Graham, J.W. (2002). Missing data: Our view of the state of the art. *PSYCHOLOGICAL METHODS*, 7, 147-177. doi: 10.1037/1082-989X.7.2.147

Te Velde, E. R. (2002). The variability of female reproductive ageing and also on how the body is built. *HUMAN REPRODUCTION UPDATE*, 8(2), 141–154. doi: 10.1093/humupd/8.2.141.

Tearne, J. E., Robinson, M., Jacoby, P., Allen, K. L., Cunningham, N. K., Li, J., & McLean, N. J. (2016). Older maternal age is associated with depression, anxiety, and stress symptoms in young adult female offspring. *JOURNAL OF ABNORMAL PSYCHOLOGY*, 125, 1-10. doi: 10.1037/abn0000119.

Tearne, J. E., Robinson, M., Jacoby, P., Li, J., Newnham, J., & McLean, N. (2015). Does late childbearing increase the risk for behavioural problems in children? A longitudinal cohort study. *PAEDIATRIC AND PERINATAL EPIDEMIOLOGY,* 29, 41-49. doi: 10.1111/ppe.12165.

Tiemeier, H., Velders, F. P., Szekely, E., Roza, S. J., Dieleman, G., Jaddoe, V. W., . . . 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 AND ADOLESCENT PSYCHIATRY,* 51, 1119-1135.e7. doi: 10.1016/j.jaac.2012.08.021.

Trillingsgaard, T., & Sommer, D. (2018). Associations between older maternal age, use of sanctions, and children's socio-emotional development through 7, 11, and 15 years. *EUROPEAN JOURNAL OF DEVELOPMENTAL PSYCHOLOGY,* 15, 141-155. doi: 10.1080/17405629.2016.1266248

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *JOURNAL OF STATISTICAL SOFTWARE,* 45, 1-67. 10.18637/jss.v045.i03.

Van Buuren, S. (2012). *FLEXIBLE IMPUTATION OF MISSING DATA.* Boca Raton: Chapman and Hall/CRC.

Weiser, M., Reichenberg, A., Werbeloff, N., Kleinhaus, K., Lubin, G., Shmushkevitch, M., . . . Davidson, M. (2008). Advanced parental age at birth is associated with poorer social functioning in adolescent males: Shedding light on a core symptom of schizophrenia and autism. *SCHIZOPHRENIA BULLETIN*, 34, 1042–1046. doi: 10.1093/schbul/sbn109

Wong, S. W., Solomon, B. D., Bodian, D. L., Kothiya, P., Eley, G., Huddleston, K. C., . . . Niederhuber, J. E. (2016). New observations on maternal age effect on germline de novo mutations. *NATURE COMMUNICATIONS,* 7, 10486. doi: 10.1038/ncomms10486.

Chapter 4 Supplement

Parental Age and Offspring Childhood Mental Health:

A Multi-Cohort, Population-Based Investigation

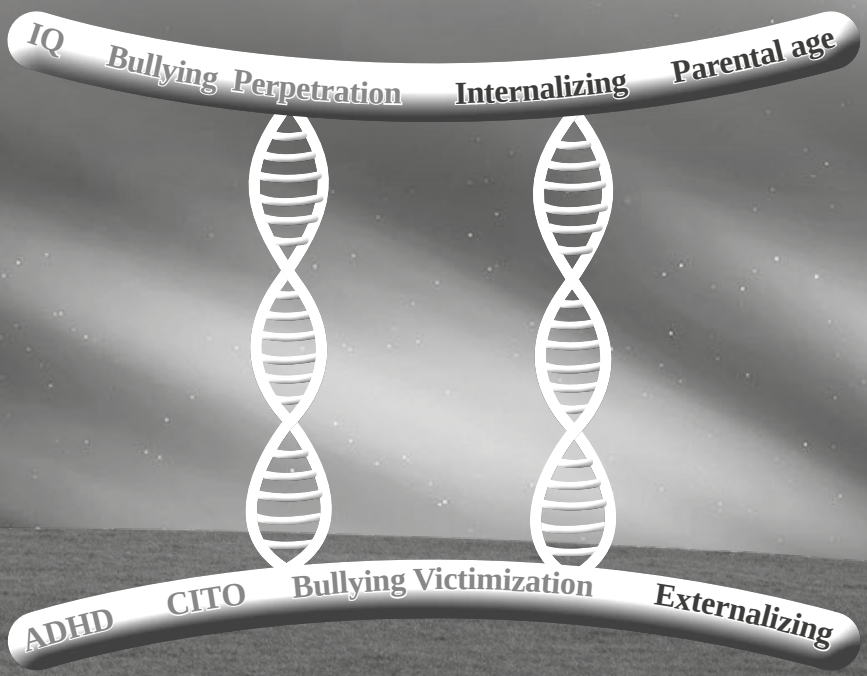


Table S1. *Mean and SD for Externalizing and Internalizing Problems of Girls*

Rater	Cohort	Externalizing	Internalizing	N-Ext/N-Int
Child	**Gen-R**	1.77 (1.81)	2.35 (2.15)	2,054/2,058
	RADAR-Y	10.19 (6.75)	0.15 (0.93)	213/120
	TRAILS	7.55 (5.40)	12.06 (7.51)	1,115/1,106
Mother	**Gen-R**	3.34 (4.29)	4.90 (4.96)	2,305/2,305
	NTR	4.79 (5.33)	4.81 (5.14)	5,626/5,577
	TRAILS	7.26 (6.21)	7.89 (6.25)	1,006/1,002
Father	**Gen-R**	3.38 (4.20)	4.54 (4.67)	1,655/1,656
	NTR	4.03 (4.84)	3.62 (4.25)	3,764/3,734
Teacher	**NTR**	2.12 (4.28)	4.29 (4.85)	3,314/3,268
	TRAILS	0.26 (0.59)	0.96 (1.09)	992/993

Table S2. *Mean and SD for Externalizing and Internalizing Problems of Boys*

Rater	Cohort	Externalizing	Internalizing	N-Ext/N-Int
Child	**Gen-R**	2.12 (2.02)	1.95 (2.00)	1,955/1,959
	RADAR-Y	10.93 (7.44)	-0.20 (0.77)	278/146
	TRAILS	9.85 (6.83)	10.47 (7.23)	1,073/1,065
Mother	**Gen-R**	4.51 (5.41)	4.81 (5.14)	2,244/2,245
	NTR	6.46 (6.73)	4.55 (4.99)	5,460/5,409
	TRAILS	9.59 (7.62)	7.80 (6.16)	959/953
Father	**Gen-R**	4.61 (5.48)	4.63 (4.77)	1,604/1,603
	NTR	5.32 (5.87)	3.50 (4.22)	3,656/3,640
Teacher	**NTR**	4.48 (6.96)	4.52 (5.08)	3,222/3,178
	TRAILS	0.63 (0.88)	1.03 (1.14)	933/931

Exploratory Results

These results are based on standardized variables, with age being mean-centered first.

Table S3. *Parental Age Predicting Externalizing Problems from Exploratory Results*

Rater	Cohort	Age F. β (p-value)	Age² F. β (p-value)	r^2	Age M. β (p-value)	Age² M. β (p-value)	r^2
Child	**Gen-R**	-.06 (<.001)	.08 (<.001)	.01	-.05 (.02)	.05 (.03)	.01
	RADAR-Y	-.05 (.44)	.14 (.05)	.02	-.08 (.22)	.18 (.01)	.04
	TRAILS	-.01 (.83)	-.01 (.77)	.00	-.03 (.39)	-.03 (.36)	.00
Mother	**Gen-R**	-.10 (<.001)	.09 (<.001)	.01	-.10 (<.001)	.02 (<.001)	.02
	NTR	-.12 (<.001)	.08 (<.001)	.01	-.11 (<.001)	.06 (<.001)	.02
	TRAILS	.09 (.02)	.08 (.04)	.01	-.13 (<.001)	.06 (.06)	.02
Father	**Gen-R**	-.10 (<.001)	.08 (.003)	.01	-.08 (.001)	.07 (<.001)	.01
	NTR	-.13 (<.001)	.07 (<.001)	.02	-.12 (<.001)	.06 (<.001)	.02
Teacher	**NTR**	-.05 (<.001)	.03 (.047)	.00	-.04 (.001)	.04 (.009)	.00
	TRAILS	-.08 (.03)	.06 (.11)	.01	-.11 (<.001)	.04 (.20)	.01

Note. F. = Father. M. = Mother.

Table S4. *Age Father and Covariates Predicting Externalizing Problems from Exploratory Results*

Rater	Cohort	Age β (p-value)	Age² β (p-value)	SES β (p-value)	Gender Child β (p-value)	r^2
Child	**Gen-R**	-.05 (.03)	.07 (<.001)	-.07 (<.001)	-.08 (<.001)	.02
	RADAR-Y	-.05 (.50)	.13 (.07)	-.06 (.39)	-.06 (.39)	.02
	TRAILS	-.01 (.88)	-.01 (.77)	-.01 (.67)	-.18 (<.001)	.03
Mother	**Gen-R**	-.09 (<.001)	.07 (.004)	-.08 (<.001)	-.15 (<.001)	.04
	NTR	-.10 (<.001)	.07 (<.001)	-.08 (<.001)	-.13 (<.001)	.05
	TRAILS	-.04 (.27)	.06 (.10)	-.17 (<.001)	-.16 (<.001)	.06
Father	**Gen-R**	-.10 (<.001)	.06 (.01)	-.06 (.03)	-.15 (<.001)	.04
	NTR	-.11 (<.001)	.06 (<.001)	-.13 (<.001)	-.14 (<.001)	.05
Teacher	**NTR**	-.04 (.006)	.02 (.125)	-.10 (<.001)	-.17 (<.001)	.04
	TRAILS	-.05 (.20)	.05 (.19)	-.13 (<.001)	-.25 (<.001)	.09

Table S5. *Age Mother and Covariates Predicting Externalizing Problems from Exploratory Results*

Rater	Cohort	**Age** ***β* (*p*-value)**	**Age²** ***β* (*p*-value)**	**SES** ***β* (*p*-value)**	**Gender Child** ***β* (*p*-value)**	r^2
Child	**Gen-R**	-.04 (.11)	.04 (.12)	-.06 (.009)	-.08 (<.001)	.02
	RADAR-Y	-.07 (.25)	.17 (.01)	.-05 (.43)	-.05 (.41)	.04
	TRAILS	-.02 (.51)	-.02 (.58)	-.01 (.79)	-.18 (<.001)	.03
Mother	**Gen-R**	-.08 (<.001)	.06 (.004)	-.06 (.006)	-.14 (<.001)	.04
	NTR	-.09 (<.001)	.06 (<.001)	-.12 (<.001)	-.14 (<.001)	.05
	TRAILS	-.08 (.02)	.06 (.06)	-.15 (<.001)	-.16 (<.001)	.07
Father	**Gen-R**	-.07 (.009)	.06 (.02)	-.05 (.09)	-.15 (<.001)	.04
	NTR	-.10 (<.001)	.05 (<.001)	-.12 (<.001)	-.14 (<.001)	.05
Teacher	**NTR**	-.03 (.035)	.03 (.019)	-.10 (<.001)	-.17 (<.001)	.04
	TRAILS	-.07 (.03)	.05 (.11)	-.12 (<.001)	-.25 (<.001)	.09

Table S6. *Exploratory Results for Parental Age Predicting Internalizing Problems*

Rater	Cohort	**Age F.** ***β* (*p*-value)**	**Age² F.** ***β* (*p*-value)**	r^2	**Age M.** ***β* (*p*-value)**	**Age² M.** ***β* (*p*-value)**	r^2
Child	**Gen-R**	-.03 (.001)	.05 (.020)	.00	-.02 (.32)	.04 (.07)	.00
	RADAR-Y	-.03 (.69)	.03 (.76)	.01	-.04 (.64)	.06 (.41)	.01
	TRAILS	.00 (.98)	-.01 (.78)	.00	-.02 (.55)	.03 (.40)	.00
Mother	**Gen-R**	-.04 (.12)	.06 (.02)	.00	-.06 (.01)	.05 (.05)	.01
	NTR	-.06 (<.001)	.05 (<.001)	.00	-.06 (<.001)	.03 (.022)	.00
	TRAILS	.01 (.81)	.05 (.17)	.00	-.05 (.12)	.04 (.26)	.00
Father	**Gen-R**	-.05 (.06)	.06 (.02)	.00	-.03 (.21)	.03 (.28)	.00
	NTR	-.07 (<.001)	.04 (.013)	.01	-.07 (<.001)	.02 (.116)	.01
Teacher	**NTR**	-.01 (.538)	.02 (.301)	.00	-.01 (.719)	.01 (.299)	.00
	TRAILS	-.02 (.56)	.01 (.89)	.00	-.04 (.21)	.04 (.20)	.00

Note. F. = Father. M. = Mother.

Table S7. *Age Father and Covariates Predicting Internalizing Problems from Exploratory Results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	-.02 (.47)	.04 (.08)	-.04 (.07)	.10 (<.001)	.02
	RADAR-Y	-.03 (.75)	.05 (.53)	-.05 (.61)	.27 (.001)	.09
	TRAILS	.01 (.84)	-.01 (.72)	-.02 (.53)	.11 (<.001)	.01
Mother	**Gen-R**	-.02 (<.001)	.03 (.20)	-.10 (<.001)	.00 (.90)	.01
	NTR	-.05 (<.001)	.04 (.001)	-.06 (<.001)	.02 (.081)	.01
	TRAILS	.03 (.48)	.04 (.24)	-.06 (.06)	.04 (.25)	.01
Father	**Gen-R**	-.04 (.14)	.04 (.09)	-.05 (.04)	-.02 (.44)	.01
	NTR	-.06 (<.001)	.03 (.034)	-.07 (<.001)	-.01 (.495)	.01
Teacher	**NTR**	-.00 (.846)	.01 (.386)	-.055 (<.001)	-.03 (.007)	.00
	TRAILS	.02 (.49)	-.02 (.54)	-.16 (<.001)	-.01 (.78)	.03

Table S8. *Age Mother and Covariates Predicting Internalizing Problems from Exploratory Results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	-.01 (.64)	.04 (.11)	-.04 (.09)	.11 (<.001)	.02
	RADAR-Y	-.02 (.78)	.09 (.23)	-.04 (.62)	.27 (.001)	.09
	TRAILS	-.02 (.63)	-.03 (.27)	-.01 (.66)	.11 (<.001)	.01
Mother	**Gen-R**	-.03 (.18)	.03 (.25)	-.09 (<.001)	.00 (.88)	.01
	NTR	-.04 (<.001)	.02 (.049)	-.06 (<.001)	.02 (.085)	.01
	TRAILS	-.04 (.30)	.03 (.34)	-.04 (.19)	.03 (.28)	.01
Father	**Gen-R**	-.02 (.56)	.02 (.52)	-.06 (.03)	.02 (.47)	.01
	NTR	-.05 (<.001)	.02 (.220)	-.07 (<.001)	-.01 (.489)	.01
Teacher	**NTR**	.00 (.936)	.01 (.434)	-.05 (<.001)	-.03 (.009)	.00
	TRAILS	.01 (.72)	.03 (.33)	-.15 (<.001)	-.01 (.71)	.03

Confirmatory Results

These results are based on unstandardized variables, only the age variables are mean-centered.

Table S9. *Exploratory Results for Parental Age Predicting Internalizing Problems*

Rater	Cohort	Age F. β (*p*-value)	Age² F. β (*p*-value)	r^2	Age M. β (*p*-value)	Age² M. β (*p*-value)	r^2
Child	**Gen-R**	-.03 (.002)	.00 (.07)	.01	-.02 (.07)	.00 (.01)	.01
	RADAR-Y	-.04 (.67)	.04 (.003)	.04	-.12 (.22)	.04 (.01)	.03
	TRAILS	-.00 (.41)	-.00 (.39)	.00	-.00 (.21)	.00 (.27)	.00
Mother	**Gen-R**	-.03 (.19)	.00 (.06)	.00	.00 (.90)	.01 (.004)	.00
	NTR	-.12 (<.001)	.01 (<.001)	.01	-.15 (<.001)	.01 (.09)	.01
	TRAILS	-.00 (.01)	.00 (.61)	.01	-.01 (<.001)	.00 (.39)	.02
Father	**Gen-R**	-.06 (.02)	.01 (.05)	.01	-.04 (.15)	.01 (.03)	.01
	NTR	-.09 (<.001)	.00 (.01)	.01	-.12 (<.001)	.00 (.18)	.01
Teacher	**NTR**	-.05 (.003)	.00 (.13)	.00	-.07 (<.001)	.01 (.06)	.00
	TRAILS	-.01 (.02)	.00 (.06)	.01	-.01 (<.001)	.00 (.12)	.02

Note. F. = Father. M. = Mother.

Table S10. *Age Mother and Covariates Predicting Externalizing Problems from Exploratory Results*

Rater	Cohort	Age β (*p*-value)	Age² β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	-.02 (.02)	.00 (.21)	-.14 (.002)	-.33 (<.001)	.02
	RADAR-Y	-.04 (.67)	.03 (.01)	1.33 (.42)	-.47 (.61)	.04
	TRAILS	-.00 (.51)	-.01 (.37)	-.01 (.12)	.04 (<.001)	.04
Mother	**Gen-R**	-.01 (.66)	.00 (.24)	-.41 (<.001)	-.96 (<.001)	.02
	NTR	-.10 (<.001)	.01 (.003)	-.84 (<.001)	-.20 (<.001)	.05
	TRAILS	-.00 (.09)	-.001 (.89)	-.04 (<.001)	.04 (<.001)	.08
Father	**Gen-R**	-.04 (.13)	.00 (.17)	-.44 (.001)	-1.08 (<.001)	.02
	NTR	-.07 (<.001)	.00 (.06)	-.75 (<.001)	-1.70 (<.001)	.04
Teacher	**NTR**	-.04 (.02)	.00 (.29)	-.52 (<.001)	-2.38 (<.001)	.05
	TRAILS	-.00 (.24)	.02 (.20)	-.08 (<.001)	.08 (<.001)	.09

Table S11. *Age Mother and Covariates Predicting Externalizing Problems from Confirmatory Results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	-.01 (.30)	.00 (.04)	-.13 (.01)	-.33 (<.001)	.02
	RADAR-Y	-.12 (.27)	.04 (.02)	1.32 (.43)	-.44 (.63)	.03
	TRAILS	-.00 (.39)	.01 (.21)	-.01 (.24)	.04 (<.001)	.04
Mother	**Gen-R**	.03 (.26)	.01 (.04)	-.43 (<.001)	-.94 (<.001)	.02
	NTR	-.12 (<.001)	.00 (.23)	-.81 (<.001)	-1.89 (<.001)	.05
	TRAILS	-.00 (.01)	.00 (.55)	-.04 (<.001)	.04 (<.001)	.08
Father	**Gen-R**	-.02 (.56)	.01 (.12)	-.43 (.003)	-1.073 (<.001)	.02
	NTR	-.09 (<.001)	.00 (.37)	-.72 (<.001)	-1.71 (<.001)	.04
Teacher	**NTR**	-.05 (.01)	.01 (.10)	-.05 (<.001)	-2.38 (<.001)	.05
	TRAILS	-.01 (.05)	.02 (.18)	-.07 (<.001)	.08 (<.001)	.09

Table S12. *Confirmatory Results for Parental Age Predicting Internalizing Problems*

Rater	Cohort	Age F. β (*p*-value)	Age^2 F. β (*p*-value)	r^2	Age M. β (*p*-value)	Age^2 M. β (*p*-value)	r^2
Child	**Gen-R**	-.01 (.22)	.00 (.25)	.00	-.01 (.44)	.00 (.05)	.00
	RADAR-Y	-.00 (.76)	.00 (.23)	.01	-.01 (.55)	.00 (.78)	.01
	TRAILS	.00 (.78)	-.00 (.88)	.00	-.00 (.71)	.00 (.38)	.00
Mother	**Gen-R**	-.05 (.02)	.02 (.03)	.00	-.04 (.08)	.01 (.01)	.01
	NTR	-.04 (.01)	.00 (.04)	.00	-.05 (.002)	.00 (.65)	.00
	TRAILS	-.00 (.91)	.00 (.83)	.00	-.00 (.17)	.00 (.40)	.00
Father	**Gen-R**	-.06 (.01)	.00 (.08)	.01	-.05 (.05)	.01 (.07)	.01
	NTR	-.02 (.18)	.00 (.29)	.00	-.02 (.21)	.00 (.55)	.00
Teacher	**NTR**	.00 (.88)	.00 (.25)	.00	.00 (.80)	.00 (.16)	.00
	TRAILS	-.00 (.40)	.03 (.03)	.01	-.01 (.06)	.03 (.01)	.01

Note. F. = Father. M. = Mother.

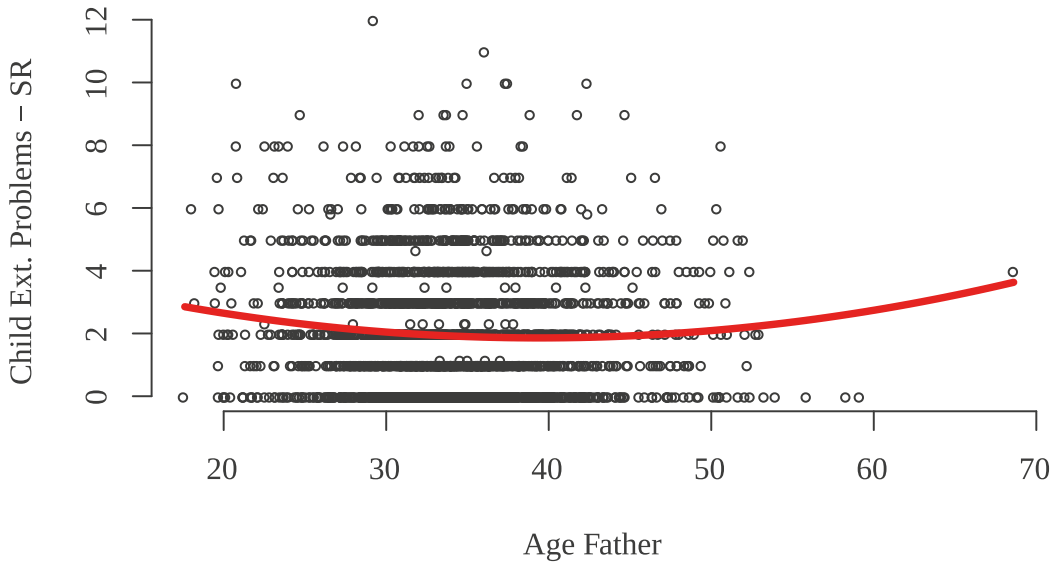
Table S13. *Age Father and Covariates Predicting Internalizing Problems from Confirmatory Results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	Gen-R	-.01 (.18)	.00 (.23)	.03 (.52)	.31 (.001)	.01
	RADAR-Y	.00 (.94)	.00 (.27)	.22 (.46)	.30 (.03)	.06
	TRAILS	.00 (.51)	-.00 (.69)	-.01 (.16)	-.04 (<.001)	.02
Mother	Gen-R	-.03 (.12)	.00 (.19)	-.47 (<.001)	.10 (.64)	.01
	NTR	-.03 (.03)	.00 (.001)	-.22 (.001)	.05 (.61)	.00
	TRAILS	.00 (.75)	-.00 (.95)	-.02 (.02)	.01 (.47)	.01
Father	Gen-R	-.04 (.06)	.00 (.32)	-.45 (.001)	-.01 (.95)	.01
	NTR	-.02 (.26)	.00 (.34)	-.08 (.24)	-.05 (.63)	.00
Teacher	NTR	.01 (.61)	.00 (.24)	-.17 (.01)	-.23 (.05)	.00
	TRAILS	.00 (.99)	.02 (.11)	-.05 (<.001)	.02 (.17)	.03

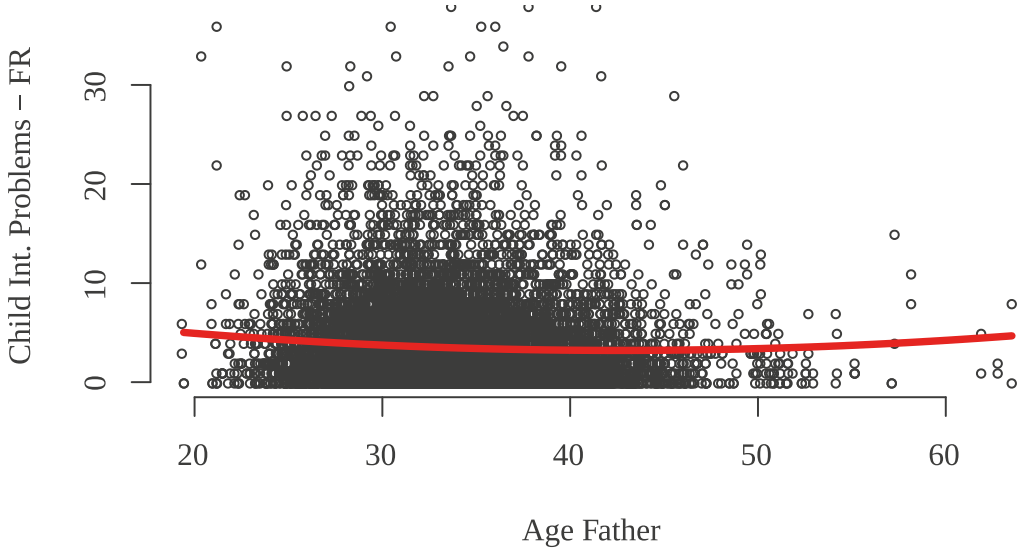
Table S14. *Age Mother and Covariates Predicting Internalizing Problems from Confimatory Results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	Gen-R	-.01 (.36)	.00 (.03)	.05 (.34)	.31 (.001)	.01
	RADAR-Y	-.00 (.84)	.00 (.73)	.24 (.42)	.29 (.03)	.05
	TRAILS	.00 (.98)	.01 (.54)	-.01 (.26)	-.03 (<.001)	.02
Mother	Gen-R	-.01 (.62)	.01 (.08)	-.46 (<.001)	.11 (.62)	.01
	NTR	-.04 (.02)	.00 (.84)	-.22 (.001)	.05 (.65)	.00
	TRAILS	-.00 (.49)	.01 (.49)	-.01 (.05)	.01 (.44)	.01
Father	Gen-R	-.02 (.38)	.01 (.23)	-.44 (.001)	-.02 (.95)	.01
	NTR	-.01 (.31)	.00 (.60)	-.07 (.26)	-.05 (.62)	.00
Teacher	NTR	.01 (.50)	.00 (.21)	-.17 (.01)	-.23 (.05)	.00
	TRAILS	-.00 (.63)	.03 (.02)	-.05 (<.001)	.02 (.15)	.03

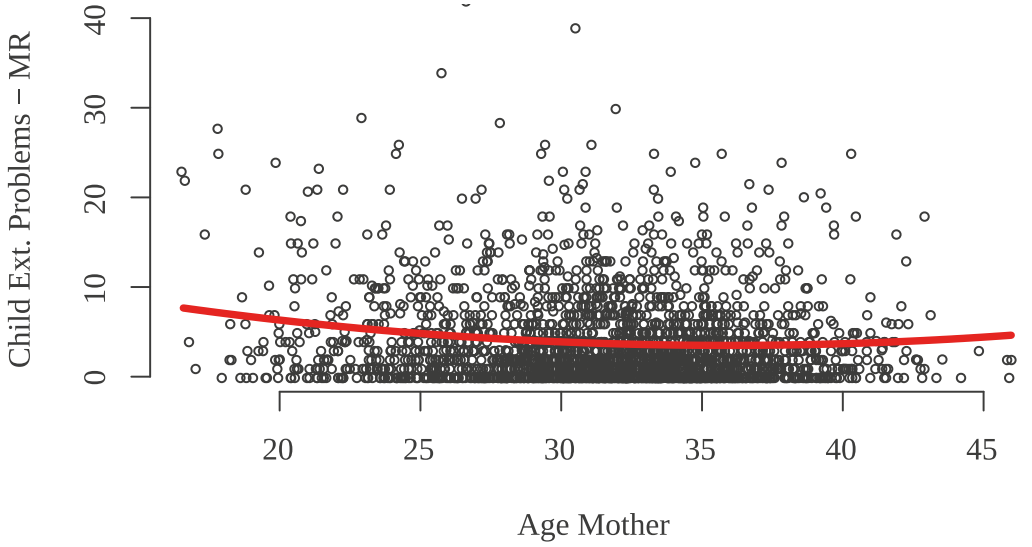
Figure S1. *Exploratory results for parental age in relation to problem behavior as represented in Gen-R and NTR.*



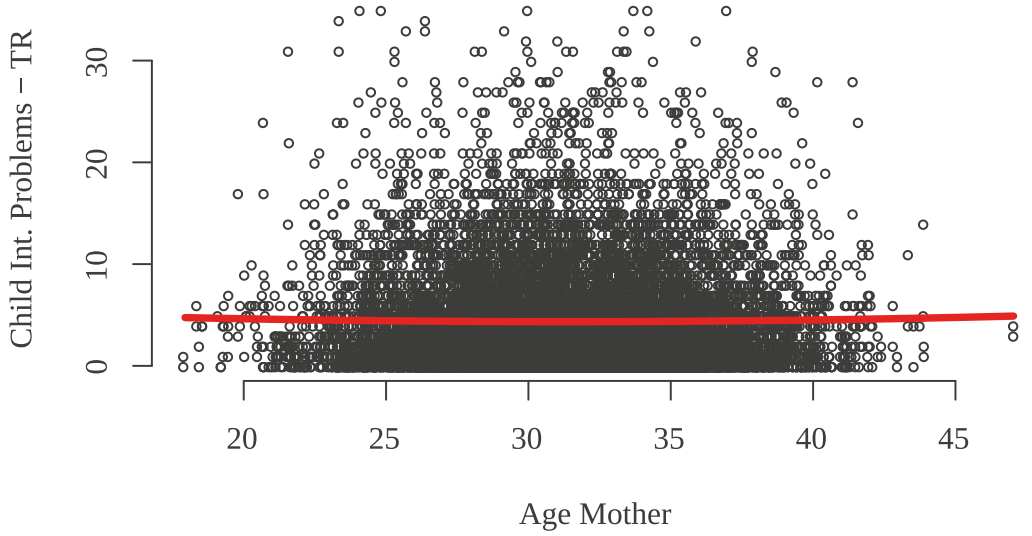
(a) Gen-R child-reported externalizing problems in relation to paternal age



(b) NTR father-reported internalizing problems in relation to paternal age



(c) Gen-R mother-reported externalizing problems in relation to maternal age



(d) NTR teacher-reported internalizing problems in relation to maternal age

The Importance of Using Cross Validation

As will be illustrated in this section, even if the size of the exploratory part of a data set is large, it may contain features that are not replicated in the confirmatory part of the data set, that is, it may contain features that are specific to the data set and not to the population from which the data are sampled.

Consider, for example, in Table S3, the exploratory results for maternal age predicting exploratory problem behavior as rated by the mother in NTR. As can be seen, the linear and quadratic effects of age are significant, that is, $p < .001$ in both cases. This relationship is also clearly visible in Figure S2. Without a cross validation approach, without much hesitation, most researchers would conclude that both the linear and quadratic effects of age exist.

Now consider, in Table 5, the corresponding confirmatory analyses. As can be seen, for NTR, with a posterior model probability of .97 H_2 is supported most by the confirmatory part of the data. As a reminder, H_2 states that there is a decreasing linear effect of maternal age on externalizing problem behavior rated by the mother, but not a quadratic effect.

There is a contrast between the results of the exploratory and confirmatory analyses. The explanation is obtained if Figure S2, constructed using the exploratory part of the data, is compared with Figure S3, constructed using the confirmatory part of the data. As can be seen, there are relatively few data points for the younger mothers. Of these data points (when randomly dividing the data set in an exploratory and a confirmatory part) relatively many corresponding to mothers giving high ratings of the externalizing problems of their children were assigned to the exploratory part of the data. Therefore, as can also be seen from the curvature of the regression line, the quadratic effect is stronger in the exploratory part than in the confirmatory part. Stated otherwise, the quadratic effect is "detected" in one part but not in the other, which implies that it is not a property of the population from which the data were sampled. Note that, similar phenomena can be observed if other corresponding pairs of exploratory and confirmatory analyses are compared.

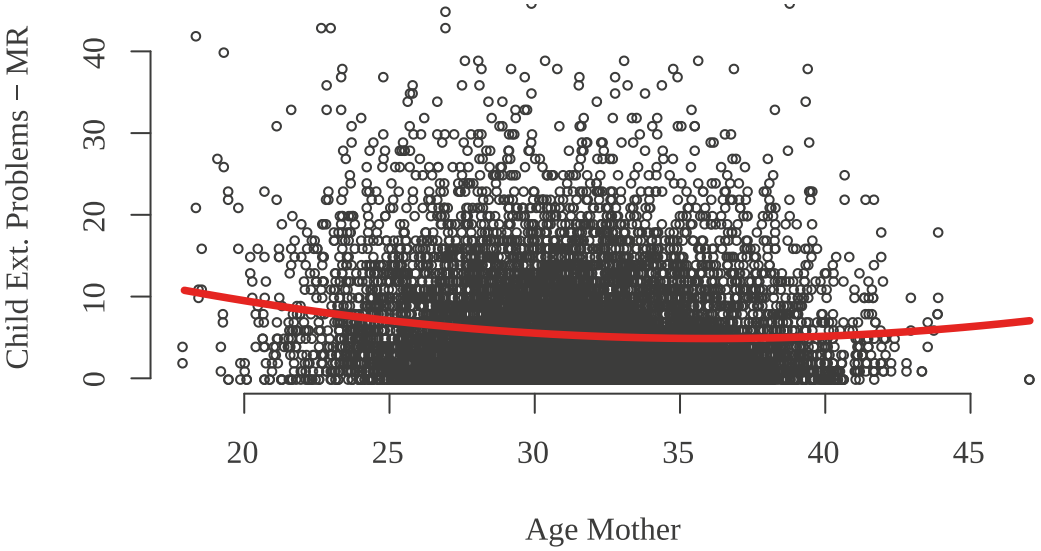


Figure S2. *NTR mother-reported externalizing problems in relation to parental age. Exploratory results.*

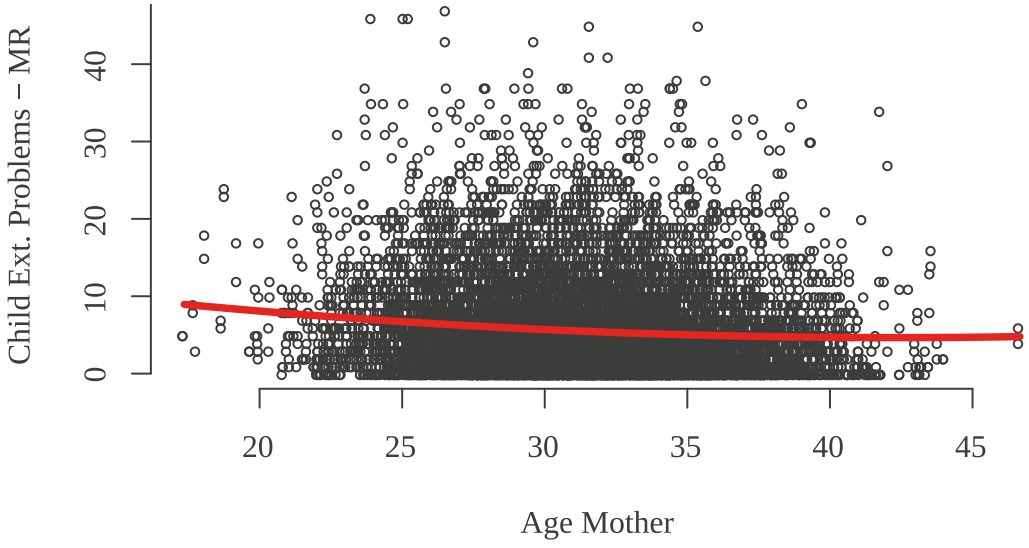
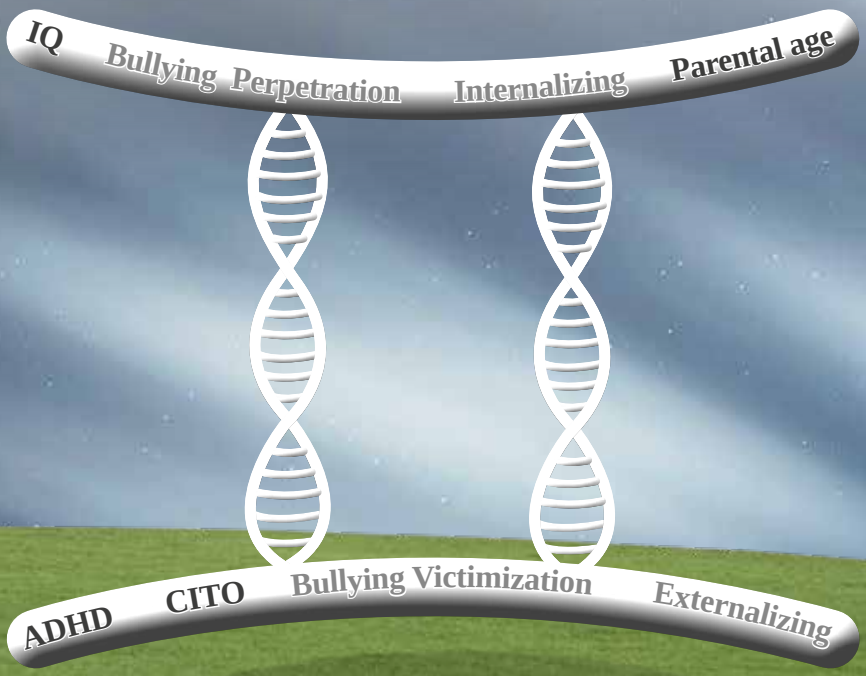


Figure S3. *NTR mother-reported externalizing problems in relation to parental age. Confirmatory results.*

Chapter 5

Effect of Parental Age on Offspring's Neurodevelopment



This chapter is based on:

Veldkamp, S.A.M.*, Zondervan-Zwijnenburg, M.A.J. *, van Bergen,, E., Barzeva, S.A. , Tamayo Martinez, N., Becht, A.I., van Beijsterveldt, C.E.M., Meeus, W., Branje, S., Hillegers, M.H.J., Oldehinkel, A.J., Hoijtink, H.J.A., Boomsma, D.I.*, & Hartman, C.* (2019). Effect of parental age on offspring's neurodevelopment. Submitted.

***These authors contributed equally to this work**

Abstract

Background. The aim of this study was to investigate whether established detrimental effects of advanced parenthood on neurodevelopmental disorders extend to the more common neurodevelopmental outcomes: attention problems, intelligence, and educational achievement. **Methods.** We analyzed child-, father-, mother- and teacher-rated attention-problems (*N*=38,024), intelligence (*N*=10,273) and educational achievement (*N*=17,522) of children from four Dutch population-based cohorts. We used 50% of the datasets to generate hypotheses and the other 50% to evaluate support for these hypotheses. With Bayesian evidence synthesis, we combined the results over cohorts. Data were analyzed with and without inclusion of child gender and SES as covariates. **Results.** We mostly found linear relations between parental age and attention-problems as reported by fathers, mothers, and teachers. Offspring of younger parents were more disadvantaged. Maternal age was also positively and linearly related to IQ and educational achievement in offspring. Paternal age showed an inverted U-shaped relation with IQ, with younger and older fathers being disadvantaged, and an attenuating positive relation with educational achievement. After including SES, we mostly found no effects. **Conclusions.** There were hardly any harmful effects of advanced parental age on attention problems, intelligence, and educational achievement. SES had an important role in the relation between parental age and offspring neurodevelopmental outcomes.

Introduction

Postponing parenthood to advanced age has been a persistent trend in all European as well as in many other developed countries during the past decades. In the Netherlands, for example, women nowadays first give birth around age 30, while in 1970 the mean age was 24 (CBS, 2019). Concerns about this postponement are understandable and growing, as a large body of research has shown that offspring of older parents are at increased risk for developing severe neurodevelopmental disorders, such as schizophrenia, Down syndrome, and autism (Merikangas, 2016; 2017). One important question is whether these effects generalize to the more common neurodevelopmental outcomes. In a recent population-based study, we found no negative effects of advanced parenthood on internalizing and externalizing problems, but observed that children of older parents tended to show fewer externalizing behavior problems than children of younger parents (Zondervan et al., in press). In the current study, we focused on attention problems and cognition and investigated whether offspring of older parents are at increased risk for more attention problems, and lower intelligence and educational achievement.

While the risk of high parental age on offspring schizophrenia, Down syndrome, and autism seems well-established, no consistent pattern exists for attention problems. Attention problems are an important component of Attention Deficit Hyperactivity Disorder (ADHD), one of the most common neurodevelopmental disorders in childhood (Faraone, Sergeant, Gillberg & Biederman, 2003). There are studies that show a reverse association, suggesting that offspring of younger parents are more at risk. Mikkelsen and colleagues (2016) found in a population-based sample (*N*=943,785) that offspring of mothers who gave birth to children early in their reproductive lives were more vulnerable to develop ADHD. This same outcome was also observed in a case-control (*N*=10,409; *N*=39,125) study by Chudal et al. (2015) and in population-based cohort studies (*N*=1,495,543; *N*=1,490,745) by Chang et al. (2014) and Janecka et al. (2019). The results are more diverse for fathers. While Mikkelsen et al. (2016) found no effect for fathers, D'Onofrio et al. (2014) reported, in a population-based study (*N*=2,615,081), that offspring of fathers 45 years and older were at higher risk for ADHD. Chudal et al. (2015), however, found that the relationship between paternal age and offspring ADHD showed high risk for young fathers (<25), lowest risk for fathers around 30, and a somewhat increased risk for fathers older than 40. Taken together, most studies point to an adverse linear effect of paternal age, but a curvilinear effect with adverse scores in both extremes of the age distribution has also been reported.

The relation between parental age and attention problems might thus differ for fathers and mothers and might also differ from those found in research on more extreme neurodevelopmental problems, such that offspring of younger parents could also be more at risk. It is therefore important to investigate not only linear effects, but curvilinear effects as well.

For intelligence, earlier studies showed mixed results. Saha et al. (2009) found in a sample of 33,437 children that intelligence at age 7 was lower for offspring of older fathers. Gajos and Beaver (2017) reported a similar finding for verbal IQ scores in daughters (N=449). McGrath, Mortensen, Pedersen, Ehrensen and Petersen (2013) found that both younger and older fathers had children with lower IQ scores than fathers aged 25-29 (i.e. an inverted U shape) (N=169,009). Gajos and Beaver (2017) reported the same effect for verbal IQ scores in sons of younger and older fathers (N=480). On the other hand, D'Onofrio et al. (2014) observed that children of fathers aged 45 or older were more vulnerable for low educational attainment and failing a grade. Regarding maternal age, some studies indicated that offspring of older mothers (and not fathers) had a higher chance of cognitive disability (Cohen, 2014), while other studies suggested that older mothers have offspring with higher IQ scores (McGrath et al., 2013). Like attention problems, effects of parental age on cognitive ability need to be further clarified.

The present study looks into the effects of parental age on neurodevelopmental outcomes. We analyzed parent-, teacher- and self-reported attention problems (N≤38,024), psychometric IQ (N=10,273) and educational achievement assessed by standardized tests (N=17,522) of children from four large population-based cohort studies. We investigated paternal age and maternal age and with and without two possible confounders: child gender and family SES. Given mixed results in previous research, we used cross-validation to generate hypotheses based on one half of our data, and subsequently evaluated how much support each of these hypotheses obtained in the other half of the data. Furthermore, Bayesian statistical methods were used to evaluate overall support.

Method

Participants

Four Dutch cohorts contributed to this study: the Netherlands Twin Register (NTR), Generation R (Gen-R), the Research on Adolescent Development and Relationships-Young cohort (RADAR-Y), and the Tracking Adolescents' Individual Lives Survey (TRAILS). The number of participants differed over dependent variables (Supplementary Tables S1-S3).

NTR recruits newborn twins from all regions in the Netherlands. Children were excluded if they had a severe handicap that interfered with daily functioning. For attention problems, we included data on 10-year-olds who were born between 1986 and 2008. The children had a mean age of 9.95 (*SD*=0.51), ranging from 7.83 to 11.95. For educational achievement, data of twins and their siblings came from a nation-wide standardized test assessed around age 12. For IQ, data of twins and their siblings measured at ages 5, 7, 9, 10, 12, 17 and 18 were included. Parental age information is given in Supplementary Table S4. Parents were mostly born in the Netherlands (95.7% of fathers and 96.7% of mothers). Mother's educational level was low (i.e., no education or primary education) for 4.6%, intermediate (i.e., secondary school, vocational training) for 67.0%, and high (i.e., bachelor's degree, university) for 28.4%.

Gen-R recruited pregnant women in the city of Rotterdam and their partners. For attention problems, 10-year-old participants were included (born between 2002 and 2006). The age of the children ranged from 8.68 to 12.47 (*M*=9.73, *SD*=0.33). For educational achievement, Gen-R analyzed data obtained from a nation-wide standardized test assessed around age 12. IQ was measured at 6 years. Information for parental age is given in Supplementary Table S4. In the overall dataset, 58.7% of the sample was of Dutch or other European ancestry, other groups included Moroccan, Dutch Antilles, and Cape-Verdian. Mother's educational level was low for 4.1%, intermediate for 39.4%, and high for 56.6%.

The RADAR-Y sample was recruited in the province of Utrecht and four large cities in the mid–west of the Netherlands. Because the RADAR-Y study had a focus on delinquency development, children with increased externalizing behavior problems at age 12 were oversampled. All participants from the first wave of data collection, born between 1990 and 1995, were selected for inclusion. The mean age was 13.03 (*SD*=0.46), ranging from 11.01-15.56. Parental age information is given in Supplementary Table S4. The sample consisted mainly of children with parents born in the Netherlands (93.3%). The other children had parents born in Surinam (1.8%), Indonesia (1%), and Dutch Antilles (0.8%). Mother's educational level was for 3.2%, intermediate for 56.7%, and high for 40.1%.

The TRAILS sample was recruited in the Northern regions of the Netherlands. All participants from the first wave of data collection (born between 1990 and 1991) were included in all analyses. Average age of the children was 11.11 (*SD*=0.56) and ranged between 10.01 and 12.58.The majority of participants had parents born in the Netherlands (86.5%), with others from Surinam (2.1%), Indonesia (1.7%), Antilleas (1.7%), Marocco (0.7%), Turkey (0.5%), and other (6.9%). Mother's educational level was low for 6.6%, intermediate for 64.3%, and high for 25.9%.

Measures

Predictors

Maternal and Paternal Age at Birth. The age of the biological parents at birth of the child was measured in years up to two decimals for each cohort (Supplementary Table S4 for descriptive statistics). Gen-R used parental age at intake (during pregnancy).

Outcomes

Attention Problems. Attention problems were measured with the ASEBA questionnaires in Gen-R, NTR, and TRAILS, which include the child-rated Youth Self Report (YSR; Achenbach & Rescorla, 2001), the parent-rated Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001, and Achenbach, 1991 for earlier birth cohorts), and the teacher-rated Teacher Report Form (TRF; Achenbach & Rescorla, 2001). Radar-Y measured mother-rated attention problems with a Dutch adaptation of Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders (DPD; Pelham, Gnagy, Greenslade & Milich, 1992; Oosterlaan, Scheres, Antrop, Roeyers, and Sergeant, 2000). In TRAILS, teachers rated child behavior on a five-point scale for: "fails to finish things he/she starts, can't concentrate, can't pay attention for long, is confused, daydreams, has learning difficulties, is clumsy or poorly coordinated, is inattentive, is easily distracted, underachieves, fails to carry out tasks". This item was derived from the set of TRF items on attention. See Supplementary Table S1 for descriptive statistics.

IQ. In Gen-R, IQ was measured using the Snijders-Oomen nonverbal intelligence test (Tellegen, Laros & Winkel, 2005). In NTR, IQ was measured using the RAKIT, WISC-R(-III), Raven or WAIS (see Franić et al., 2014). For the children in NTR with multiple assessments, the mean over all IQ assessments was taken. In Radar-Y and TRAILS, IQ was assessed with the block design and the vocabulary subtests of the WISC-III-R (Legerstee, van der Reijden-Lakeman, Lechner-van der Noort & Ferdinand, 2004). See Supplementary Table S2 for descriptive statistics.

Educational Achievement. Educational achievement was available in two cohorts: Gen-R and NTR. Scores came from a 3-day nation-wide standardized test around age 12 (end of primary school; Citogroep, 2019). Most schools in the Netherlands take part. See Supplementary Table S3 for descriptive statistics.

Covariates

Socio-Economic Status (SES) and child gender. In Gen-R, SES was defined as a continuous variable (principal component) based on parental education and household income. In NTR, SES was a 5-level ordinal variable based on parental occupational level. In TRAILS, SES was a 3-level ordinal variable based on parental education, parental occupational status and household income. In RADAR-Y SES was a dichotomous variable based on parents' occupational level. Child gender was coded as male = 0 and female = 1.

Missing Data and Data Imputation

The proportions of missing data per cohort and variable are provided in Supplementary Tables S5-S7. NTR used different subgroups for attention problems (only twins) and cognitive functioning (educational achievement and IQ data was also available for triplets and siblings). In addition, note that IQ data were only available in a subset (~10%), so IQ was analyzed with a subset of the cognitive functioning dataset, including only children for who at least one IQ assessment was available (to prevent introducing bias by imputing a large proportion of the data). Therefore, IQ has no missing values. For Gen-R, educational achievement data was only available for a small subset of the overall dataset (26.8%), therefore a sub-dataset was used containing participants with complete educational achievement data only.

Missing data were imputed (Schafer & Graham, 2002; Van Buuren, 2012) with the package mice (Van Buuren & Groothuis-Oudshoorn, 2011) in R (R Core Team, 2018). The imputation was conducted for attention and the cognitive functioning datasets separately, which all included variables on paternal age, maternal age, SES and child gender. Datasets were split into an exploratory and confirmatory half (see analytical strategy). Except for participant and family ID, all variables in the datasets were selected as predictors in the imputation model if the correlation was larger than .10 with the to be imputed variable. The data were imputed 100 times, and analyses results were pooled over these datasets by the mice package. The imputation for the twins of Gen-R and NTR was done per twin-pair instead of per participant, to ensure equal information within twins on parental age and SES. The (non-twin) sibling data was imputed as in the other cohorts.

Analytical Strategy

The analytical strategy consisted of four steps: (1) exploratory data analysis, (2) informative hypothesis generation, (3) Bayesian hypothesis evaluation in confirmatory data per cohort, and (4) Bayesian evidence synthesis over cohorts.

Exploratory Data Analysis

As previous research is mixed about the relations between parental age and the outcome variables, we started with exploratory data analyses. In each cohort, the datasets were randomly divided into an exploratory and a confirmatory part. In the exploratory data, linear regression analyses were conducted in R with as predictors standardized father age and father age squared, or mother age and mother age squared. The dependent variables were attention problems (reported by either child, father, mother, or teacher) child IQ, and educational achievement. The analyses were first conducted without covariates. Next, gender was added as a covariate, and thirdly, SES was added as a covariate. For the datasets including twins or siblings (i.e., Gen-R and NTR), data were split based on Family ID to create independent datasets (so that all siblings are in one dataset) and linear regression analyses were cluster-corrected based on Family ID with the R-package lavaan (Rosseel, 2012).

Informative Hypothesis Generation

Informative hypotheses are hypotheses that contain information about the parameters of interest in the model, like that a regression parameter is positive (Hoijtink, 2012). Based on the direction and significance of the exploratory regression analyses, competing informative hypotheses were composed stating that the βage and βage2 parameters were either negative, equal to zero, or positive. In the set of competing hypotheses, two hypotheses were included by default: the null hypothesis: βage=0, βage2=0, and the unconstrained alternative hypothesis: βage, βage2. The unconstrained alternative hypotheses entails that 'anything goes', that is: βage, βage2 can take on any value. This alternative hypothesis is a fail-safe hypothesis that will receive most support when the informative hypotheses in the set do not represent the data well.

Bayesian Hypothesis Evaluation in Confirmatory Data per Cohort

In the confirmatory data, linear regression analyses were conducted with mean-centered father or mother age and age squared as predictors, and the same dependent variables and covariates as before. Using the statistical software Bain (Gu et al., 2017; Hoijtink, Gu, & Mulder, 2018), the relative support of each informative hypothesis versus the unconstrained alternative (i.e., βage, βage2) was computed. Posterior model probabilities (PMPs) represented the relative probability of each of the evaluated hypotheses in the set, summing up to 1.00.

Bayesian Evidence Synthesis over Cohorts

Next, results were updated over cohorts, meaning that we evaluated which informative hypothesis was best supported by all cohorts simultaneously. In this step we can unite results from cohorts that used different measures, because we evaluate informative hypotheses that are applicable irrespective of the operationalization of the attention and cognitive constructs.

Assessing how much the hypotheses are supported by all cohorts, evaluates support for hypotheses irrespective of the population and measurement specifics of separate cohorts.

Results

Exploratory Data Analyses

In general, the results of the exploratory analyses indicated that child-reported attention problems were not predicted by parental age (results are provided in Supplementary Tables S8-S21). For other reporters, age had a significant negative relation with attention problems, accompanied by a significant positive quadratic factor in about half of the analyses across raters and cohorts. The negative direction of the linear relation indicated that offspring of younger parents had on average more attention problems. In case of significant quadratic factors, the regression either became U-shaped, indicating that offspring of the youngest and oldest parents had most attention problems or had a steeper decline in the beginning that attenuated over time, indicating that offspring of the youngest parents had the most attention problems (see for example Figure 1a-1b). For parental age with IQ and educational achievement the linear relations were positive: offspring of younger parents had on average lower IQ or educational achievement. Also, significant quadratic factors were now negative resulting in either a bow-shape (inverse U), indicating that offspring of the youngest and oldest parents had the lowest IQ and educational achievement scores or had a steeper increase in the beginning that attenuated over time. Offspring of the youngest parents had the lowest IQ and educational achievement (see for example Figure 1c-1d). Adding gender as a covariate to the model did generally not change the patterns. When SES was added to the model about half of the significant relations between age and attention problems disappeared.

Informative Hypothesis Generation in Exploratory Data

Based on the exploratory results, the overall set of hypotheses for attention problems was:

- H_1: $\beta_1=0$, $\beta_2=0$. Age is unrelated (i.e., the classical null model).
- H_2: $\beta_1<0$, $\beta_2=0$. Age has a negative linear relation, there is no quadratic relation.
- H_3: $\beta_1<0$, $\beta_2>0$. Age has a negative linear relation, and a positive quadratic relation.
- H_4: $\beta_1=0$, $\beta_2>0$. Age has a positive quadratic relation, there is no linear relation.
- H_a: β_1, β_2. The relation with age can be anything.

A set of these competing hypotheses was drafted for each combination of predictor (paternal age or maternal age), dependent variable (i.e., attention rated by mother, father, teacher, child; IQ; educational achievement), and set of covariates (i.e., none or gender and SES). For example, for teacher reported attention problems regressed on maternal age, we found H_3: $\beta_1<0$, $\beta_2>0$ in NTR and H_2: $\beta_1<0$, $\beta_2=0$ in TRAILS. As a fail-safe, we always evaluate H_1: $\beta_1=0$, $\beta_2=0$, and H_a: β_1, β_2 (see Analytical Strategy section). Hence, we evaluated H_1, H_2, H_3, and H_a as the set of competing hypotheses with the confirmatory data in all cohorts for the regression of teacher reported attention problems on maternal age. See Supplementary Table S22 for the exact hypotheses for attention problems per rater, before and after adjustment for gender and SES. Note that we composed hypotheses and ran analyses with gender and SES in the model at once, because gender on itself hardly affected any of the relations in the model. For IQ and educational achievement, the overall set of hypotheses was:

- H_1: $\beta_1=0$, $\beta_2=0$. Age is unrelated (i.e., the classical null model).
- H_2: $\beta_1>0$, $\beta_2=0$. Age has a positive linear relation, there is no quadratic relation.
- H_3: $\beta_1>0$, $\beta_2<0$. Age has a positive linear relation, and a negative quadratic relation.
- H^4: $\beta_1=0$, $\beta_2<0$. Age has a negative quadratic relation, there is no linear relation.
- H_a: β_1, β_2. The relation with age can be anything.

See Supplementary Table S23 for the exact hypotheses for IQ and educational achievement before and after adjustment for gender and SES.

Bayesian Hypothesis Evaluation and Evidence Synthesis in Confirmatory Data

Cohort-specific and robust results are provided in Tables 1 to 6. Cohort-specific results are fully described in the Supplementary Tables S24-S32. We focus on the robust results across cohorts.

First, for attention problems, child-reported data showed no relation with parental age across cohorts. For all other informants, results without covariates supported a negative linear relation between parental age and attention problems, i.e. fewer attention problems in offspring of older parents. One exception is that overall, there was no relation between paternal age and mother-reported attention problems. When including gender and SES in the model, we found most support for no relation between attention problems and parental age. Two exceptions were the relation between father-reported attention problems and paternal age, and mother-reported problems with maternal age. Most support was found for a negative linear relation with parental age (i.e., older parents reported less attention-problems); even after including covariates. Second, for IQ, most support was found for a quadratic relation with paternal age with slightly lower scores for younger and older fathers (inverted U; see Figure 2a-2c), or a relation that attenuated with older age (see Figure 2d).

A positive linear relation between maternal age and IQ was found. After taking child gender and SES into account, the relation with IQ disappeared for paternal age, but the linear relation was still best supported for maternal age. Third, for educational achievement, the findings of the two largest cohorts (Gen-R and NTR) indicated that there was a quadratic relation with parental age, in which children of younger fathers (see Figure 3a-3b) and younger mothers (see Figure 3c-3d) were disadvantaged. Offspring of older mothers had higher educational achievement. For both parents, the effects disappeared after taking child gender and SES into account.

Table 1. *Posterior Model Probabilities for Parental Age Predicting Attention Problems*

Rater	Cohort	Age Father					Age Mother				
		H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
Child	**Gen-R**	*1.00*	-	-	-	.01	*1.00*	-	-	-	.00
	TRAILS	*1.00*	-	-	-	.00	*1.00*	-	-	-	.00
	All	**1.00**	-	-	-	.00	**1.00**	-	-	-	.00
Mother	**Gen-R**	*.44*	-	.38	.07	.12	.30	*.44*	.21	-	.06
	NTR	.04	-	*.73*	.00	.23	.00	*.58*	.33	-	.09
	TRAILS	*.78*	-	.12	.06	.04	.01	*.78*	.17	-	.05
	RADAR-Y	*.71*	-	.12	.13	.04	.06	*.56*	.31	-	.08
	All	**.71**	-	.28	.00	.00	.00	**.97**	.03	-	.00
Father	**Gen-R**	.04	*.80*	.12	-	.04	*.56*	.32	.09	-	.03
	NTR	.09	*.77*	.11	-	.04	.01	*.84*	.12	-	.03
	All	.01	**.97**	.02	-	.00	.01	**.95**	.04	-	.00
Teacher	**NTR**	*.94*	.06	-	-	.00	*.91*	.08	.01	-	.00
	TRAILS	.02	*.95*	-	-	.04	*.00*	.41	*.47*	-	.12
	All	.25	**.75**	-	-	.00	.00	**.93**	.07	-	.01

Note. Numbers in *italic* font represent the highest posterior model probability per cohort. Numbers in **bold** font represent the highest results after Bayesian updating.

Dashes indicate that the hypothesis was not among the set of evaluated hypotheses based on the exploratory analyses.

Table 2. *Posterior Model Probabilities for Parental Age Predicting Attention Problems After Correction for Covariates*

Rater	Cohort	Age Father					Age Mother				
		H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
Child	**Gen-R**	*1.00*	-	-	-	.00	*1.00*	-	-	-	.00
	TRAILS	*1.00*	-	-	-	.00	*1.00*	-	-	-	.00
	All	**1.00**	-	-	-	.00	**1.00**	-	-	-	.00
Mother	**Gen-R**	*.85*	.09	.01	.05	.00	*.88*	.05	.00	.06	.00
	NTR	.33	*.62*	.04	.01	.01	.03	*.85*	.10	.00	.03
	TRAILS	*.91*	.04	.00	.04	.00	*.42*	*.36*	.10	.09	.03
	RADAR-Y	*.55*	.31	.05	.07	.02	.11	*.60*	.19	.05	.05
	All	**.99**	.01	.00	.00	.00	.11	**.89**	.00	.00	.00
Father	**Gen-R**	.25	*.73*	-	-	.02	*.91*	.08	-	-	.00
	NTR	*.66*	.34	-	-	.01	.43	*.57*	-	-	.01
	All	.40	**.60**	-	-	.00	**.89**	.11	-	-	.00
Teacher	**NTR**	*1.00*	-	-	-	.00	*.97*	-	-	.03	.00
	TRAILS	*.98*	-	-	-	.02	*.31*	-	-	*.38*	.31
	All	**1.00**	-	-	-	.00	**.96**	-	-	.04	.00

Note. See Table 1.

Table 3. *Posterior Model Probabilities for Parental Age Predicting IQ*

Cohort	Age Father					Age Mother				
	H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
Gen-R	.00	.00	*.77*	.00	.23	.00	*.37*	*.51*	.00	.12
NTR	*.56*	.27	.06	.10	.02	*.53*	*.30*	.06	.09	.02
TRAILS	.00	*.76*	.19	.00	.06	.00	*.62*	.31	.00	.08
RADAR-Y	*.41*	.09	.32	.13	.04	.05	.06	*.36*	*.43*	.09
All	.00	.00	**.99**	.00	.01	.00	**.51**	**.48**	.00	.00

Note. See Table 1.

Table 4. *Posterior Model Probabilities for Parental Age Predicting IQ After Correction for Covariates*

Cohort	Age Father					Age Mother				
	H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
Gen-R	.22	*.71*	.05	.01	.02	.05	*.94*	-	-	.01
NTR	*.82*	.10	.01	.07	.00	*.87*	.12	-	-	.00
TRAILS	*.65*	.29	.02	.04	.01	.02	*.94*	-	-	.04
RADAR-Y	*.51*	.10	.09	.27	.03	*.38*	*.34*	-	-	.28
All	**.97**	.03	.00	.00	.00	.01	**.99**	-	-	.00

Note. See Table 1.

Table 5. *Posterior Model Probabilities for Parental Age Predicting Educational Achievement*

Cohort	Age Father					Age Mother				
	H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
Gen-R	.00	.00	*.76*	-	.24	.00	.06	*.76*	-	.18
NTR	-	-	-	-	-	-	-	-	-	-
TRAILS	.00	.31	*.52*	-	.17	.00	*.70*	.24	-	.06
RADAR-Y	-	-	-	-	-	-	-	-	-	-
All	.00	.00	**.91**	-	.09	.00	.19	**.77**	-	.05

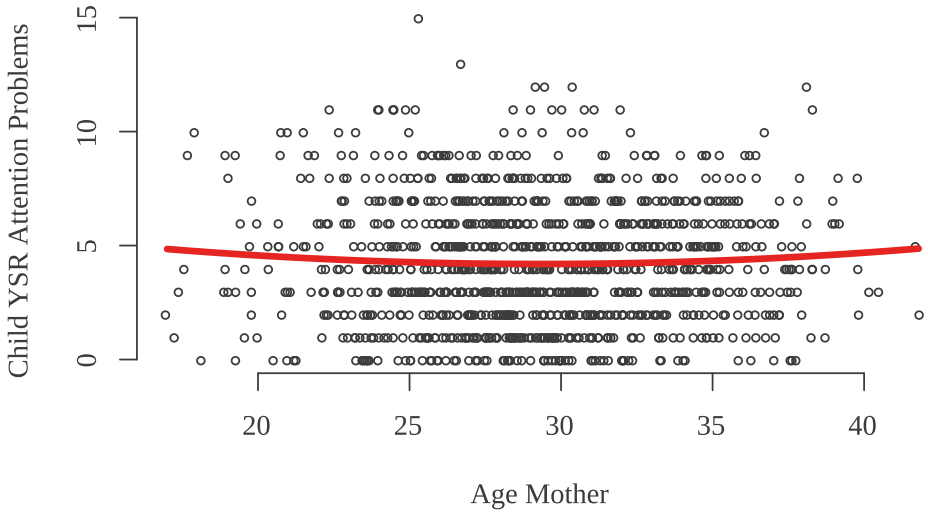
Note. See Table 1.

Table 6. *Posterior Model Probabilities for Parental Age Predicting Educational Achievement After Correction for Covariates*

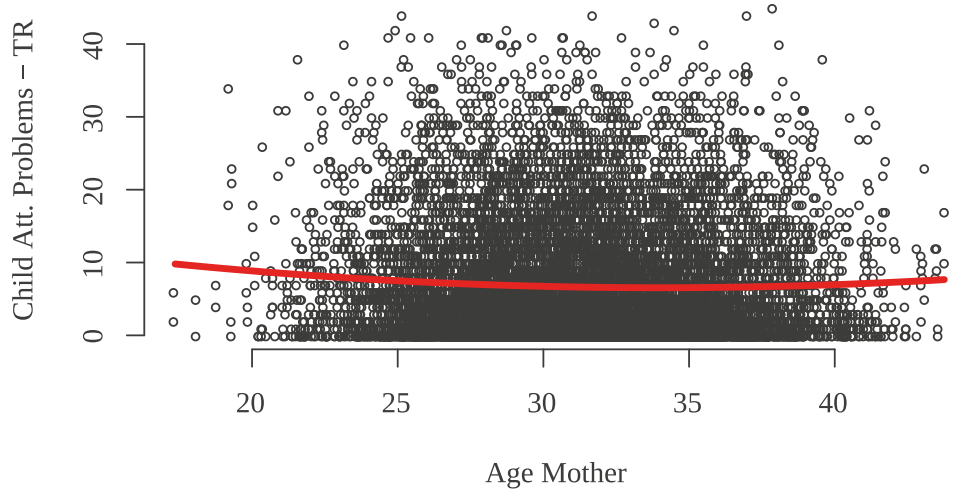
Cohort	Age Father					Age Mother				
	H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
Gen-R	*.95*	.05	-	-	.00	*.98*	.02	-	-	.00
NTR	-	-	-	-	-	-	-	-	-	-
TRAILS	*.54*	.45	-	-	.01	.09	*.89*	-	-	.02
RADAR-Y	-	-	-	-	-	-	-	-	-	-
All	**.96**	.04	-	-	.00	**.84**	.16	-	-	.00

Note. See Table 1.

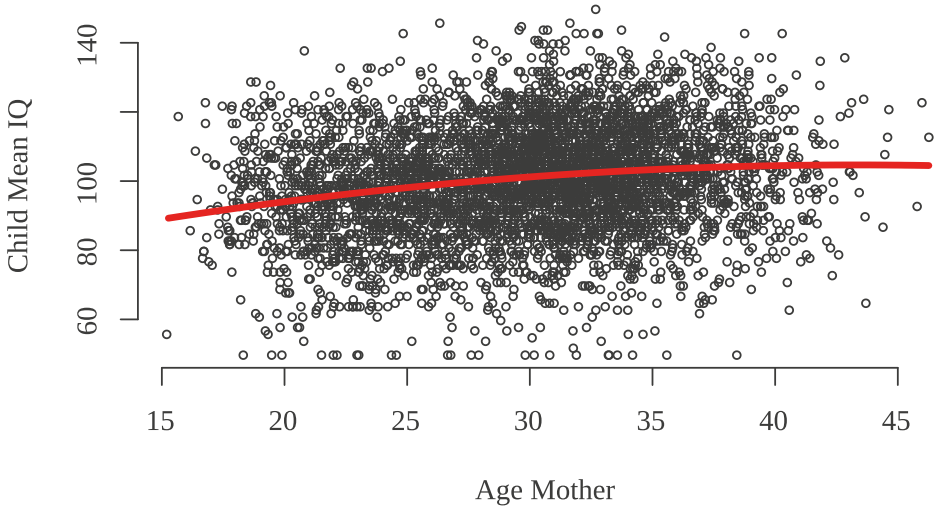
Figure 1. *Exploratory plots*



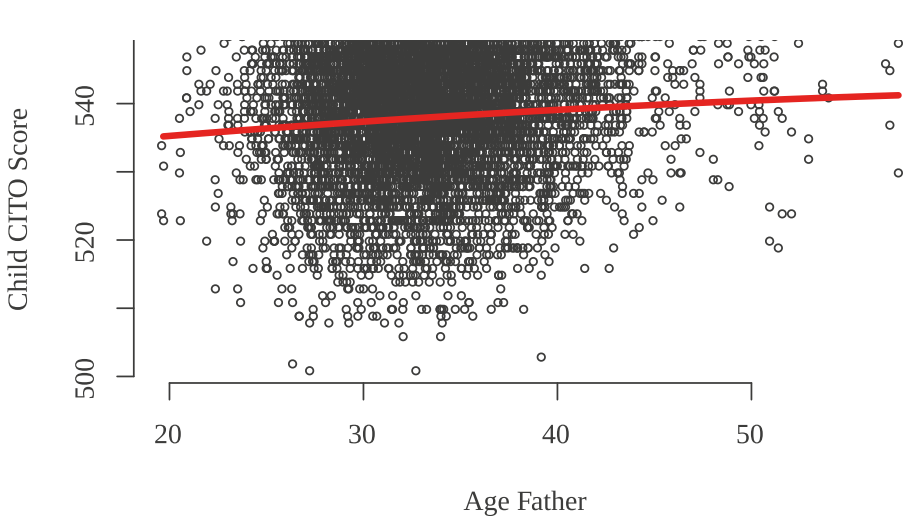
(a) Child-reported attention by TRAILS with $\beta_1 = 0$, $\beta_2 = 0$



(b) Teacher reported Attention by NTR with $\beta_1 < 0$, $\beta_2 > 0$

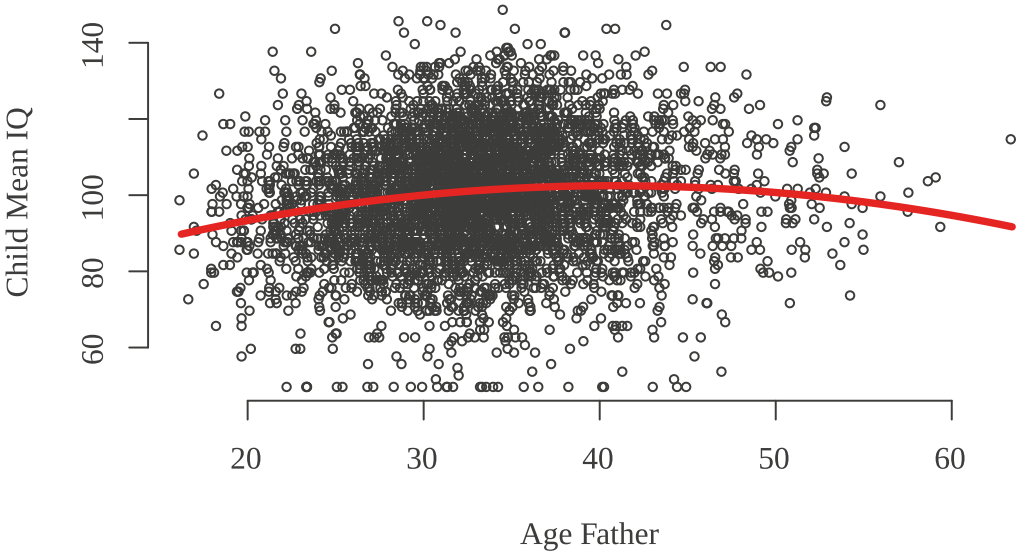


(c) IQ regressed on mother age by Gen-R with $\beta_1 > 0$, $\beta_2 < 0$

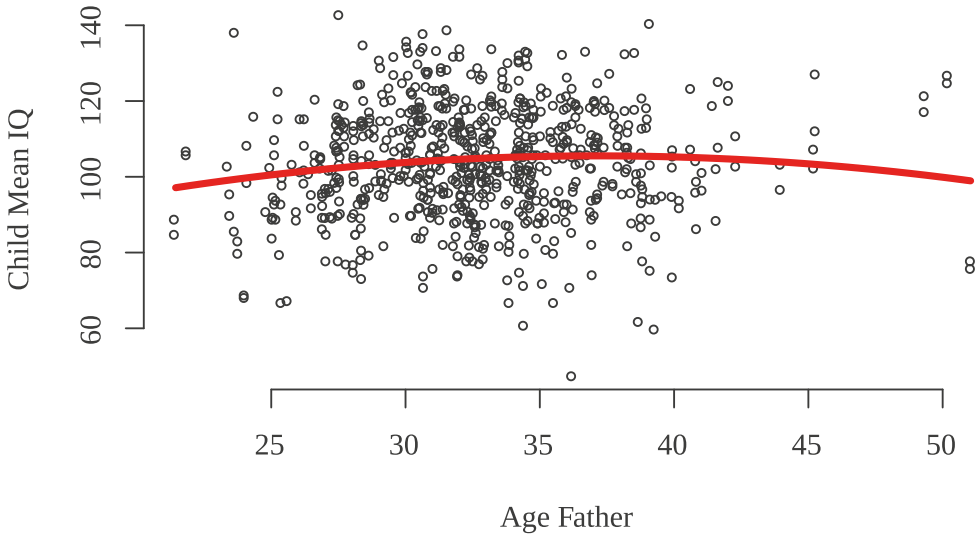


(d) Educational Achievement by NTR with $\beta_1 > 0$, $\beta_2 = 0$

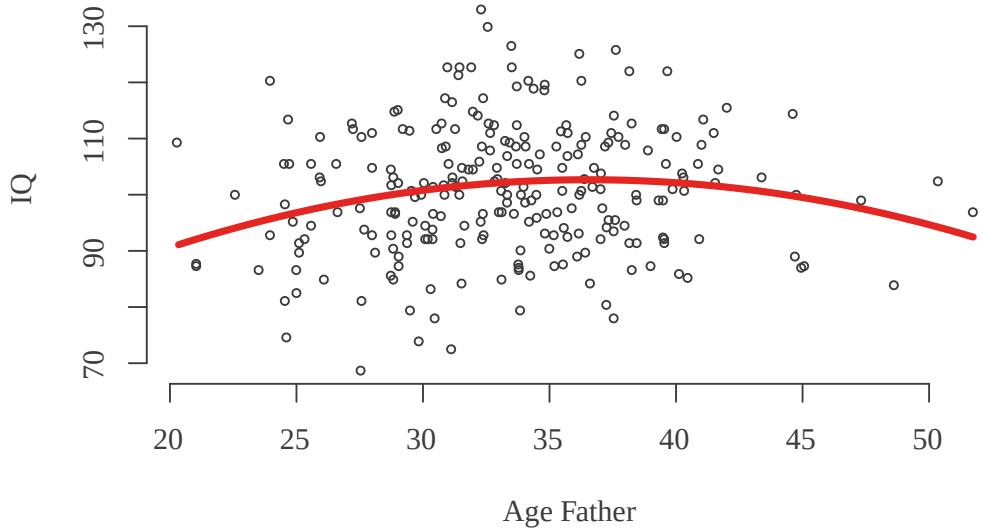
Figure 2. *Confirmatory plots for Age Father with IQ*



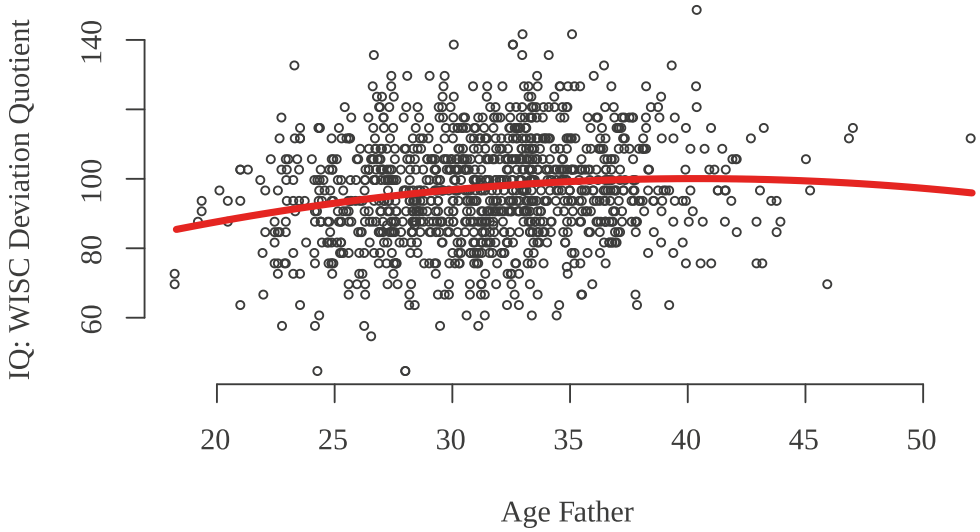
(a) Gen-R



(b) NTR

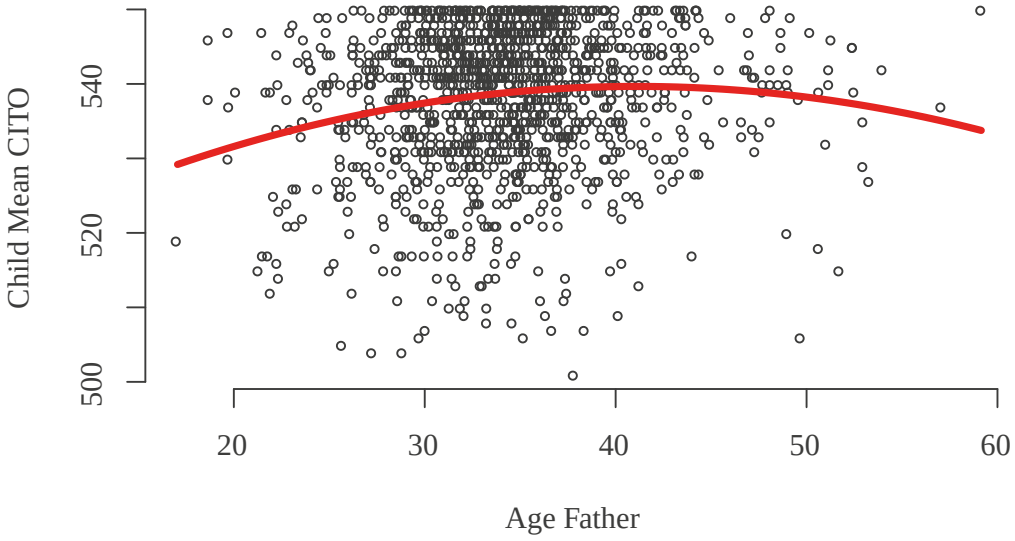


(c) Radar-Y

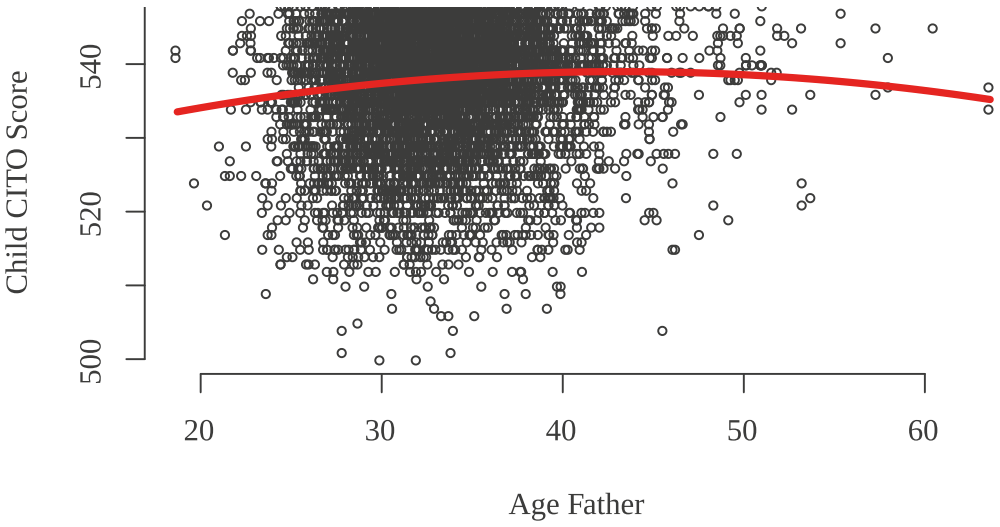


(d) TRAILS

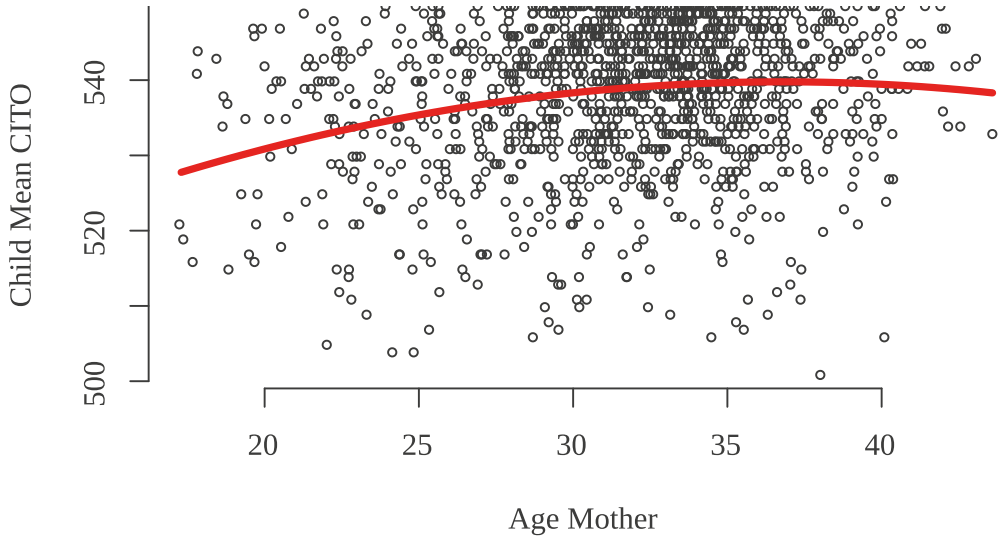
Figure 3. *Confirmatory plots for Parental Age with Educational Achievement*



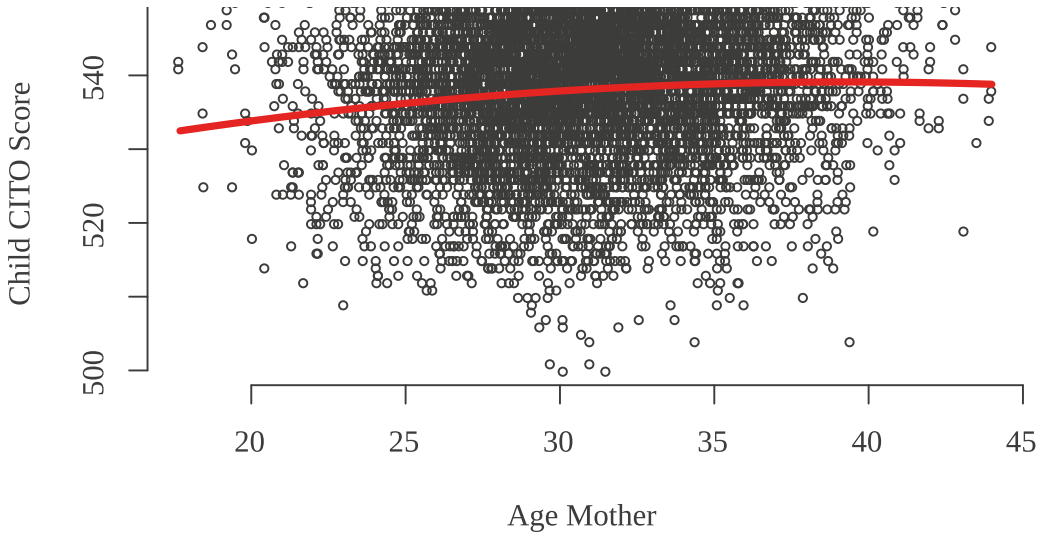
(a) Gen-R – Paternal age



(b) NTR – Paternal Age



(c) Gen-R – Maternal age



(d) NTR – Maternal Age

Discussion

We found that older parents are beneficial for offspring attention, IQ, and educational achievement. In contrast to being disadvantaged from a biological point of view (e.g. Malaspina, 2001), older parents seem to provide benefits for offspring on a psychosocial or contextual level (Janecka et al., 2019). Parents who postpone parenthood are typically highly educated with higher incomes at the time they start a family. This puts them in a better position to provide their children with a more stimulating environment (e.g., more books at home; van Bergen et al., 2017), which has been positively associated with educational attainment (Melhuish et al., 2008; Kong et al., 2018). We observed only advantageous effects of advanced parental age, and suggest that biological disadvantages appear compensated by the positive contextual factors for attention, IQ and educational achievement. This might not be the case for the more severe neurodevelopmental disorders, such as autism, where adverse effects of advanced parenthood have been found in multiple studies (reviewed by e.g. De Kluiver, Buizer-Voskamp, Dolan, & Boomsma, 2017), but is in line with the support for an advantageous relation between older age and offspring's reduced externalizing problem behavior that we found in our earlier study (Zondervan-Zwijnenburg et al., in press). The effects of parental age on child outcomes were highly similar for paternal and maternal age. This corroborates the major influence of level of SES.

Most of the statistically significant associations between parental age and child attention problems, IQ, and educational achievement disappeared when SES was taken into account. Associations that attenuate after taking SES into account suggest that most of the effect of parental age on offspring development is due to genetic and environmental effects from parent SES to child outcome. Because it is not clear which genetic and environmental effects SES captures, we argue that it is important to present results both with and without controlling for SES. Furthermore, we know that low SES tends to be associated with young parenthood, parental ADHD and lower IQ, and that low SES may reflect a more general genetic liability that influences both age at having offspring and offspring outcome. Alternatively, SES could influence parental age which, in turn, influences offspring outcome. In that case, adjusting for SES could introduce bias (Janecka et al., 2019). Hence, we conclude that older parents tend to have offspring with better attention and cognition, but the effects are small and mostly explained by higher SES.

Besides environmental transmission, parent and child characteristics are also associated due to genetic transmission. Both ADHD and intelligence are heritable phenotypes.

Individuals with ADHD and/or low IQ have an increased risk of impulsive behavior, which could result in early pregnancies (Østergaard, Dalsgaard, Faraone, Munk-Olsen & Laursen. 2017). Offspring of these young parents may thus have a genetic predisposition to develop ADHD and low IQ. Support for this hypothesis was also reported by Chung et al. (2014) and Mikkelsen et al (2016). Individuals who become parents at later ages tend to have higher educational attainment, and these parents pass on favorable education-related genetic variants. Likewise, Swagerman et al. (2017), for example, found resemblance between parents and children in reading ability was solely due to genetic transmission.

In the exploratory phase, the four cohorts consistently showed associations in the same direction (offspring of older parents performed better), but these associations did not consistently reach significance despite large samples, suggesting that the associations tended to be small. Our cross-cohort differences may relate to birth-cohorts differences. For example, Goisis (2017) found that the association between advanced maternal age and children's cognitive ability changed from negative to positive in different birth-cohorts because of changing parental characteristics. RADAR-Y and TRAILS have an early nineties cohort, and Gen-R a cohort from after 2000. Our largest cohort, NTR, included children from the 80's, 90's, and 2000's. It is unclear, however, whether there is a birth-cohort effect within this range of twenty years. Other reasons for cross-cohort differences may be structural differences between the populations, and reliability and validity of measures. Although the cohorts had some different properties and results sometimes differed, the cohorts did not yield contradictory findings. Moreover, our analytical strategy enabled us to summarize the evidence per hypothesis over cohorts.

Previous studies regarding attention problems, IQ, and educational achievement showed mixed results, but these studies used different populations, measures, covariates, etcetera. A strength of our study is that we applied Bayesian evidence synthesis, allowing us to combine evidence from multiple cohort studies that possibly used different measures. As a result, we were able to identify consistent effects and hypotheses that received the most support over cohorts. The overall results pointed towards robust effects, as they were supported by all cohorts, irrespective of the characteristics of the populations or specifics of the measurements used. Furthermore, we included large population-based samples, handled missing data by means of multiple imputation, and used cross-validation.

In conclusion, we found support for older parents having offspring with better attention, IQ and educational achievement scores and younger parents having offspring with worse attention, IQ and educational achievement scores. Only paternal age had a clear inverted U-shaped relation with educational achievement, with both offspring of younger and older fathers being disadvantaged.

More resources and more education-elevating genetic variants in older parents may compensate for possible biological disadvantages. Genetic effects in which ADHD, cognitive functioning, and young parenthood come together may explain why lower parental age goes together with more offspring problems. After including SES in the model, most of the associations with parental age disappeared. Hence, we can be certain that SES takes on an important role which may be due to genetic sharing of SES with parental age and outcomes or SES may influence parental age, which, in turn, influences offspring outcome. Based on this population-based multi-cohort study, we conclude that offspring of older parents are not disadvantaged with respect to the investigated cognitive constructs, at least where this pertains to mild outcomes as studied in the general population.

References

Achenbach, T. M. (1991). *MANUAL FOR THE CHILD BEHAVIOR CHECKLIST/4-18.* Burlington, VT: University of Vermont.

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002). Ten-year comparisons of problems and competencies for national samples of youth: Self, parent, and teacher reports. *JOURNAL OF EMOTIONAL AND BEHAVIORAL DISORDERS,* 10(4), 194-203.

Achenbach, T.M. & Rescorla, L.A. (2001). *MANUAL FOR THE ASEBA SCHOOL-AGE FORMS AND PROFILES.* Burlington, VT: University of Vermont, Research center for children, youth and families.

Centraal Bureau voor de Statistiek. (2019, January 23). *VROUWEN STEEDS LATER MOEDER*. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2018/05/vrouwen-steeds-later-moeder

Chang, Z., Lichtenstein, P., D'Onofrio, B. M., Almqvist, C., Kuja-Halkola, R., Sjölander, A., & Larsson, H. (2014). Maternal age at childbirth and risk for ADHD in offspring: A population-based cohort study. *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY,* 43(6), 1815-1824.

Chudal, R., Joelsson, P., Gyllenberg, D., Lehti, V., Leivonen, S., Hinkka-Yli-Salomäki, S., ... & Sourander, A. (2015). Parental age and the risk of attention-deficit/hyperactivity disorder: A nationwide, population-based cohort study. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY*, 54(6), 487-494.

Citogroep. (2019). Eindtoets Basisonderwijs. Retrieved from: https://www.cito.com/we-serve/primary-education

Cohen, P. N. (2014). Parental age and cognitive disability among children in the United States.

D'Onofrio, B. M., Rickert, M. E., Frans, E., Kuja-Halkola, R., Almqvist, C., Sjölander, A., ...

& Lichtenstein, P. (2014). Paternal age at childbearing and offspring psychiatric and academic morbidity. *JAMA PSYCHIATRY,* 71(4), 432-438.

De Kluiver, H., Buizer-Voskamp, J. E., Dolan, C. V., & Boomsma, D. I. (2017). Paternal age and psychiatric disorders: A review. *AMERICAN JOURNAL OF MEDICAL GENETICS PART B*: Neuropsychiatric Genetics, 174, 202–213. doi: 10.1002/ajmg.b.32508.

Faraone, S. V., Sergeant, J., Gillberg, C., & Biederman, J. (2003). The worldwide prevalence of ADHD: Is it an American condition?. *WORLD PSYCHIATRY,* 2(2), 104.

Franić, S., Dolan, C. V., van Beijsterveldt, C. E., Pol, H. E. H., Bartels, M., & Boomsma, D. I. (2014). Genetic and environmental stability of intelligence in childhood and adolescence. *TWIN RESEARCH AND HUMAN GENETICS,* 17(3), 151-163.

Gajos, J. M., & Beaver, K. M. (2017). The role of paternal age in the prediction of offspring intelligence. *THE JOURNAL OF GENETIC PSYCHOLOGY,* 178(6), 319-333.

Goisis, A., Schneider, D. C., & Myrskylä, M. (2017). The reversing association between advanced maternal age and child cognitive ability: Evidence from three UK birth cohorts. International *JOURNAL OF EPIDEMIOLOGY,* 46(3), 850-859.

Hoijtink, H. (2012). Informative hypotheses: *THEORY AND PRACTICE FOR BEHAVIORAL AND SOCIAL SCIENTISTS.* Boca Raton, FL: CRC Press.

Hvolgaard Mikkelsen, S., Olsen, J., Bech, B. H., & Obel, C. (2016). Parental age and attention-deficit/hyperactivity disorder (ADHD). *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY,* 46(2), 409-420.

Janecka, M., Hansen, S. N., Modabbernia, A., Browne, H. A., Buxbaum, J. D., Schendel, D. E., ... & Grice, D. E. (2019). Parental age and differential estimates of risk for neuropsychiatric disorders: Findings from the Danish birth cohort. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY.*

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., ... & Gudbjartsson, D. F. (2018). The nature of nurture: Effects of parental genotypes. *SCIENCE, 359*(6374), 424-428. doi: 10.1126/science.aan6877

Legerstee, J, van der Reijden-Lakeman, I.A., Lechner-van der Noort, M. G., Ferdinand, R.F. (2004). Bruikbaarheid verkorte versie WISC-RN in de kinderpsychiatrie [Usability shortened version WISC-RN in child psychiatry]. *KIND EN ADOLESCENT,* 25, 178-182.

Malaspina, D. (2001). Paternal factors and schizophrenia risk: de novo mutations and imprinting. *SCHIZOPHRENIA BULLETIN,* 27(3), 379-393.

McGrath, J., Mortensen, P. B., Pedersen, C. B., Ehrenstein, V., & Petersen, L. (2013). Paternal age and general cognitive ability-a cross sectional study of Danish male conscripts. *PLOS ONE,* 8(10), e77444.

Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *JOURNAL OF SOCIAL ISSUES,* 64(1), 95-114.

Merikangas, A. K., Calkins, M. E., Bilker, W. B., Moore, T. M., Gur, R. C., & Gur, R. E. (2017). Parental age and offspring psychopathology in the Philadelphia Neurodevelopmental Cohort. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY,* 56(5), 391-400.

Merikangas, A. K., Segurado, R., Kelleher, E., Hogan, D., Delaney, C., Gill, M., ... & Heron, E. A. (2016). Parental age, birth order and neurodevelopmental disorders. *MOLECULAR PSYCHIATRY,* 21(6), 728-730.

Oosterlaan, J., Scheres, A., Antrop, I., Roeyers, H., & Sergeant, J. A. (2000). Vragenlijst voor Gedragsproblemen bij Kinderen (VvGK) [Questionnaire for Behavioral Problems in Children 6-16 years]. Nederlandse bewerking van de Disruptive Behavior Disorders Rating Scale [Dutch translation of the Disruptive Behavior Disorders Rating Scale]. *LISSE: SWETS ZEITLINGER.*

Østergaard, S. D., Dalsgaard, S., Faraone, S. V., Munk-Olsen, T., & Laursen, T. M. (2017). Teenage parenthood and birth rates for individuals with and without attention-deficit/hyperactivity disorder: A nationwide cohort study. Journal of the American Academy of Child & *ADOLESCENT PSYCHIATRY,* 56(7), 578-584.

Pelham Jr, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY,* 31(2), 210-218.

R Core Team. (2018). A language and environment for statistical computing. Vienna, Austria: Foundation for Statistical Computing.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *JOURNAL OF STATISTICAL SOFTWARE*, 48(2), 1-36.

Saha, S., Barnett, A. G., Foldi, C., Burne, T. H., Eyles, D. W., Buka, S. L., & McGrath, J. J. (2009). Advanced paternal age is associated with impaired neurocognitive outcomes during infancy and childhood. *PLOS MEDICINE,* 6(3), e1000040.

Schafer, J.L. & Graham, J.W. (2002). Missing data: Our view of the state of the art. *PSYCHOLOGICAL METHODS,* 7, 147-177. doi: 10.1037/1082-989X.7.2.147

Silverstein, A. B. (1972). Validity of WISC-R short forms. *JOURNAL OF CLINICAL PSYCHOLOGY*, 31, 696-697.

Swagerman, S. C., Van Bergen, E., Dolan, C., de Geus, E. J., Koenis, M. M., Pol, H. E. H., & Boomsma, D. I. (2017). Genetic transmission of reading ability. *BRAIN AND LANGUAGE,* 172, 3-8.

Tellegen, P. J., Winkel, M., Wijnberg-Williams, B. J., & Laros, J. A. (2005). Snijders-Oomen niet-verbale intelligentietest SON-R 2, 5-7. Amsterdam: Boom Testuitgevers.

van Bergen, E., van Zuijen, T., Bishop, D., & de Jong, P. F. (2017). Why are home literacy environment and children's reading skills associated? What parental skills reveal. *READING RESEARCH QUARTERLY,* 52(2), 147-160.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *JOURNAL OF STATISTICAL SOFTWARE*, 45, 1-67. 10.18637/jss.v045.i03.

Van Buuren, S. (2012). *FLEXIBLE IMPUTATION OF MISSING DATA.* Boca Raton: Chapman and Hall/CRC

Chapter 5 Supplement

Effect of Parental Age on Offspring's Neurodevelopment

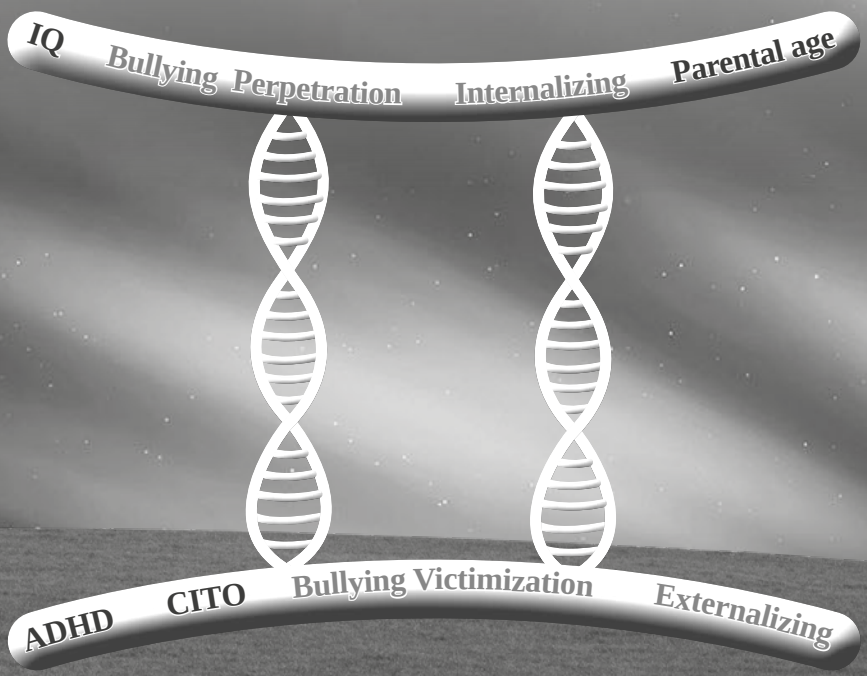


Table S1. *Mean, SD and Sample Size for Attention Problems*

Variable	Gen-R (*N* = 9,901)		NTR (*N* = 25,396)		RADAR-Y (*N* = 497)		TRAILS (*N* = 2,230)	
Informant	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)	*N*	Mean (SD)	*N*
Attention Problems								
Child	3.41 (2.49)	4,357	-	-	-	-	4.33 (2.74)	2,197
Mother	3.25 (3.20)	4,920	2.95 (3.05)	22,045	8.94[1] (8.37)	489	4.36 (3.47)	1,964
Father	3.29 (3.08)	3,555	2.62 (2.88)	14,725	-	-	-	-
Teacher	-	-	6.74 (7.87)	12,573	-	-	0.53[2] (0.58)	1,927

Note. The total sample size is presented between brackets. The sample size for each outcome variable is presented below to provide insight in the amount of missing values. Unless otherwise specified, Gen-R, NTR and TRAILS used the ASEBA questionnaires (YSR, CBCL, and TRF) to measure attention problems (Achenbach, 1991; Achenbach & Rescorla, 2001).
[1] Radar-Y measured mother-rated attention problems with a Dutch adaptation of Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders (DPD; Pelham, Gnagy, Greenslade & Milich, 1992), by Oosterlaan, Scheres, Antrop, Roeyers, and Sergeant, (2000).
[2] TRAILS uses a 1-item adapted version of the TRF (scale and range = 0-2), see Measures section for more information.

Table S2. *Mean, SD and Sample Size for IQ*

Cohort	Mean (SD)	*N*
Gen-R (*N*=6,111)	100.7[1] (15.18)	6,111
NTR (*N*=1,495)	103.44[2] (14.21)	1,495
TRAILS (*N*=2,230	97.19 (15.00)	2,221
RADAR (*N*=497)	102.05 (11.80)	446

Note. The total sample size is presented between brackets. The sample size for each outcome varia ble is presented below to provide insight in the amount of missing values. In TRAILS and Radar-Y, IQ was assessed with the block design and the vocabulary subtests of the WISC-III-R.
[1] Snijders-Oomen nonverbal intelligence test (Tellegen, Laros & Winkel, 2005)
[2] IQ was measured at ages 5, 7, 9, 10, 12, 17 and 18 using the RAKIT, WISC-R(-III), Raven or WAIS (see Franić et al., 2014). For the children in NTR with multiple assessments, the mean over all IQ assessments was taken

Table S3. *Mean, SD and Sample Size for IQ*

Cohort	Mean (SD)	*N*
Gen-R (*N*=2,655)	538.4 (9.44)	2,655
NTR (*N*=15,693)	538 (8.55)	14,867
TRAILS	-	-
RADAR	-	-

Note. The total sample size is presented between brackets. The sample size for each outcome variable is presented below to provide insight in the amount of missing values. Educational achievement was assessed by the CITO End of Primary Education Test

Table S4. *Parental Age at Offspring Birth*

Variable	Maternal age at birth child		Paternal age at birth child	
	Range	*M (SD)*	Range	*M (SD)*
Attention Problems				
Gen-R	15.27- 46.34	29.92 (5.37)	14.87- 68.18	32.99 (6.01)
NTR	17.36-47.09	31.35 (3.95)	18.75-63.61	33.76 (4.71)
RADAR-Y	17.80-48.61	31.38 (4.43)	20.34-52.52	33.70 (5.10)
TRAILS	16.34-44.88	29.32 (4.58)	18.28-52.09	32.00 (4.71)
IQ				
Gen-R	15.27-34.14	29.92 (5.37)	14.87- 68.18	32.99 (6.01)
NTR	19.26-45.63	30.18 (3.81)	19.68-57.00	32.54 (4.45)
RADAR-Y	17.80-48.61	31.38 (4.43)	20.34-52.52	33.70 (5.10)
TRAILS	16.34-44.88	29.32 (4.58)	18.28-52.09	31.995 (4.71)
Educational Achievement				
Gen-R	16.85- 46.34	32.25 (4.71)	16.48- 68.18	33.92 (5.5)
NTR	17.15-45.63	31.02 (3.80)	18.71-63.61	33.40 (4.52)
RADAR-Y	-	-	-	-
TRAILS	-	-	-	-

Note. Gen-R and NTR had different datasets for attention problems, IQ, and EA, therefore all descriptive statistics for parental age are given, since these are key variables in our study.

Imputation: Missing Data (N + %) Per Dependent Variable Dataset

Table S5. *Attention Problems*

	Child Age *N(%)*	Maternal Age *N(%)*	Paternal Age *N(%)*	Gender *N(%)*	SES *N(%)*
Gen-R*	31 (0.6)	0 (0)	645 (13.11)	0 (0)	1020 (20.73)
NTR	3,233 (12.73)	76 (.30)	332 (1.31)	0 (0)	751 (2.96)
RADAR-Y	0 (0)	2 (.40)	48 (9.66)	0 (0)	8 (1.61)
TRAILS	0 (0)	114 (5.11)	558 (25.02)	0 (0)	42 (1.88)

* This are the missing in the dataset that have mother's attention (N= 4920)

	Mother Attention *N(%)*	Father Attention *N(%)*	Teacher Attention *N(%)*	Child Attention *N(%)*
Gen-R*	0	1,527 (31,04)	-	756 (15.37)
NTR	3,351 (13.19)	10,671 (42.02)	12,782 (50.33)	-
RADAR-Y	8 (1.61)	-	-	-
TRAILS	266 (11.93)	-	303 (13.59)	33 (1.48)

* This are the missing in the dataset that have mother's attention (N= 4920)

Table S6. *CITO / Educational Achievement*

	Child Age *N(%)*	Maternal Age *N(%)*	Paternal Age *N(%)*	Gender *N(%)*	SES *N(%)*	CITO / SP *N(%)*	IQ *N(%)*
Gen-R*	0 (0)	0 (0)	372 (14.01)	0 (0)	575 (21.66)	0(0)	375 (14.12)
NTR	Exact Age is not available. Groep 8	26 (.17)	136 (.87)	0	1,046 (6.67)	826 (5.26)	14,198 (90.47)
RADAR-Y	-	-	-	-	-	-	-
TRAILS	-	-	-	-	-	-	-

* This are the missing in the dataset that have information on CITO (N= 2655)

Table S7. *IQ*

	Maternal Age *N(%)*	Paternal Age *N(%)*	Gender *N(%)*	SES *N(%)*	CITO / SP *N(%)*	IQ *N(%)*
Gen-R*	0 (0)	1,104 (18.07)	0 (0)	1,672 (27.36)	3,831 (62.69)	0 (0)
NTR	1 (.07)	11 (.74)	0	30 (2.00)	-	
RADAR-Y	2 (.40)	48 (9.66)	0 (0)	8 (1.61)	7 (1.41)	51 (10.26)
TRAILS	-	-	-	-	-	-

* This are the missing in the dataset that have information on IQ (N= 6111)

Additional Remarks/Information

- **NTR (CITO + IQ):** For families with more than one twin-pair, the 'second' twin-pairs were excluded.
- **NTR (IQ):** there were no missings, since children were selected to have IQ present due to large proportion of missing + CITO was not included in the imputation (since there were no missings for IQ)
- **Radar:** Separate imputations (with varying aux variables) were conducted for attention and IQ/Educational Achievement, but the imputation was all on the same dataset (no subset selections), because the largest percentage of missing data is ≈ 10%.
- **GenR:** The datasets were created and imputed separately (with the same aux variables) for attention, IQ and CITO, due to large proportion of missings.
- **TRAILS:** Same as Radar: Separate imputations (with varying aux variables) were conducted for attention and IQ, but the imputation was all on the same dataset (no subset selections).

Tables for Exploratory Data

Attention Problems

Table S8. *Parental Age Predicting Attention Problems from Exploratory Results*

Rater	Cohort	Age F. β (*p*-value)	Age^2 F. β (*p*-value)	r^2	Age M. β (*p*-value)	Age^2 M. β (*p*-value)	r^2
Child	**Gen-R**	.00 (.77)	.03 (.08)	.001	-.01 (.72)	.02 (.16)	.001
	TRAILS	.02 (.59)	-.01 (.81)	.001	-.00 (.95)	.03 (.30)	.001
Mother	**Gen-R**	-.05 (.003)	.03 (.03)	.002	-.03 (.03)	.04 (.01)	.003
	NTR	-.05 (<.001)	.02 (.04)	.002	-.05 (<.001)	.02 (.13)	.002
	TRAILS	-.07 (.07)	.08 (.03)	.009	-.14 (<.001)	.05 (.16)	.021
	RADAR-Y	-.12 (.114)	.20 (.006)	.035	-.13 (.041)	.20 (.002)	.045
Father	**Gen-R**	-.06 (<.001)	.03 (.046)	.003	-.03 (.03)	.03 (.04)	.002
	NTR	-.05 (<.001)	.02 (.11)	.003	-.05 (<.001)	.01 (.20)	.002
Teacher	**NTR**	-.04 (.002)	.01 (.40)	.001	-.03 (.01)	.02 (.03)	.001
	TRAILS	-.09 (.014)	.07 (.068)	.010	-.11 (<.001)	.02 (.527)	.013

Note. F. = Father. M. = Mother.

Table S9. *Age Father and Gender Child Predicting Attention Problems from Exploratory Results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	.01 (.76)	.03 (.07)	-.09 (<.001)	.009
	TRAILS	.02 (.581)	-.01 (.82)	.03 (.460)	.001
Mother	**Gen-R**	-.05 (.002)	.03 (.02)	-.14 (<.001)	.022
	NTR	-.06 (<.001)	.03 (.02)	-.15 (<.001)	.024
	TRAILS	-.07 (.055)	.08 (.036)	-.16 (<.001)	.034
	RADAR-Y	-.12 (.100)	.20 (.007)	-.11 (.075)	.047
Father	**Gen-R**	-.06 (<.001)	.03 (.04)	-.13 (<.001)	.02
	NTR	-.06 (<.001)	.02 (.08)	-.15 (<.001)	.024
Teacher	**NTR**	-.04 (<.001)	.01 (.25)	-.26 (<.001)	.066
	TRAILS	-.09 (.009)	.06 (.076)	-.17 (<.001)	.038

Table S10. *Age Father and Covariates Predicting Attention Problems from Exploratory Results*

Rater	Cohort	**Age** ***β*** **(*p*-value)**	**Age²** ***β*** **(*p*-value)**	**SES** ***β*** **(*p*-value)**	**Gender Child** ***β*** **(*p*-value)**	r^2
Child	**Gen-R**	.02 (.14)	.01 (.39)	-.07 (<.001)	-.09 (<.001)	.013
	TRAILS	.02 (.613)	-.01 (.294)	.01 (.871)	.02 (.463)	.001
Mother	**Gen-R**	-.01 (.07)	.01 (.56)	-.12 (<.001)	-.14 (<.001)	.035
	NTR	-.04 (<.001)	.02 (.11)	-.09 (<.001)	-.15 (<.001)	.033
	TRAILS	-.03 (.490)	.05 (.156)	-.20 (<.001)	-.15 (<.001)	.072
	RADAR-Y	-.12 (.111)	.20 (.007)	.04 (.545)	-.12 (.067)	.052
Father	**Gen-R**	-.04 (.03)	.01 (.36)	-.08 (<.001)	-.13 (<.001)	.027
	NTR	-.04 (<.001)	.01 (.24)	-.08 (<.001)	-.15 (<.001)	.030
Teacher	**NTR**	-.02 (.08)	.00 (.88)	-.14 (<.001)	-.26 (<.001)	.085
	TRAILS	-.05 (.186)	.04 (.294)	-.21 (<.001)	-.16 (<.001)	.079

Table S11. *Age Mother and Gender Child Predicting Attention Problems from Exploratory Results*

Rater	Cohort	**Age** ***β*** **(*p*-value)**	**Age²** ***β*** **(*p*-value)**	**Gender Child** ***β*** **(*p*-value)**	r^2
Child	**Gen-R**	-.01 (.66)	.02 (.18)	-.09 (<.001)	.009
	TRAILS	-.00 (.946)	.03 (.310)	.02 (.489)	.002
Mother	**Gen-R**	-.03 (.02)	.04 (.01)	-.14 (<.001)	.022
	NTR	-.05 (<.001)	.02 (.06)	-.15 (<.001)	.025
	TRAILS	-.14 (<.001)	.05 (.109)	-.16 (<.001)	.047
	RADAR-Y	-.14 (.038)	.19 (.004)	-.10 (.135)	.054
Father	**Gen-R**	-.03 (.02)	.03 (.045)	-.13 (<.001)	.02
	NTR	-.05 (<.001)	.02 (.10)	-.15 (<.001)	.023
Teacher	**NTR**	-.04 (<.001)	.03 (.004)	-.26 (<.001)	.067
	TRAILS	-.11 (<.001)	.03 (.408)	-.17 (<.001)	.041

Table S12. *Age Mother and Covariates Predicting Attention Problems from Exploratory results*

Rater	Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	.02 (.15)	.01 (.47)	-.07 (<.001)	-.09 (<.001)	.013
	TRAILS	-.01 (.849)	.03 (.271)	.01 (.665)	.02 (.496)	.002
Mother	**Gen-R**	.02 (.22)	.02 (.16)	-.13 (<.001)	-.14 (<.001)	.036
	NTR	-.03 (.002)	.02 (.16)	-.09 (<.001)	-.15 (<.001)	.033
	TRAILS	-.08 (.017)	.04 (.271)	-.18 (<.001)	-.16 (<.001)	.076
	RADAR-Y	-.13 (.045)	.19 (.004)	.03 (.590)	-.10 (.123)	.056
Father	**Gen-R**	.00 (.83)	.02 (.23)	-.09 (<.001)	-.13 (<.001)	.026
	NTR	-.03 (.002)	.01 (.23)	-.08 (<.001)	-.15 (<.001)	.030
Teacher	**NTR**	-.01 (.34)	.02 (.03)	-.14 (<.001)	-.26 (<.001)	.085
	TRAILS	-.05 (.118)	.01 (.807)	-.20 (<.001)	-.16 (<.001)	.078

IQ

Table S13. *Parental Age Predicting IQ from Exploratory Results*

Cohort	Age F. β (*p*-value)	Age^2 F. β (*p*-value)	r^2	Age M. β (*p*-value)	Age^2 M. β (*p*-value)	r^2
Gen-R	.14 (<.001)	-.09 (<.001)	.02	.18 (<.001)	-.06 (<.001)	.039
TRAILS	.17 (<.001)	-.12 (<.001)	.033	.21 (<.001)	-.05 (.130)	.045
NTR	.02 (.68)	-.08 (.29)	.005	-.00 (.95)	-.09 (.12)	.009
RADAR-Y	.14 (.060)	-.02 (.781)	.019	.11 (.114)	-.08 (.316)	.015

Table S14. *Age Father and Gender Child Predicting IQ from Exploratory Results*

Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	Gender Child β (*p*-value)	r^2
Gen-R	.14 (<.001)	-.09 (<.001)	-.00 (.88)	.02
TRAILS	.17 (<.001)	-.12 (<.001)	-.10 (.001)	.042
NTR	.02 (.68)	-.08 (.29)	.00 (.95)	.005
RADAR-Y	.14 (.067)	-.02 (.757)	-.09 (.167)	.028

Table S15. *Age Father and Covariates Predicting IQ from Exploratory Results*

Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Gen-R	.05 (.001)	-.02 (.16)	.37 (<.001)	-.00 (.83)	.146
TRAILS	.09 (.005)	-.08 (.017)	.35 (<.001)	-.11 (<.001)	.146
NTR	-.04 (.44)	-.04 (.52)	.24 (<.001)	-.00 (.97)	.062
RADAR-Y	.13 (.078)	-.02 (.753)	-.07 (.312)	-.08 (.207)	.033

Table S16. *Age Mother and Gender Child Predicting IQ from Exploratory Results*

Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	Gender Child β (*p*-value)	r^2
Gen-R	.18 (<.001)	-.06 (<.001)	-.00 (.94)	.039
TRAILS	.21 (<.001)	-.04 (.159)	-.10 (.001)	.054
NTR	-.00 (.95)	-.09 (.12)	-.00 (.92)	.009
RADAR-Y	.11 (.119)	-.09 (.115)	-.09 (.167)	.027

Table S17. *Age Mother and Covariates Predicting IQ from Exploratory Results*

Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Gen-R	.05 (<.001)	-.01 (.33)	.36 (<.001)	-.00 (.36)	.146
TRAILS	.10 (<.001)	-.01 (.704)	.34 (<.001)	-.10 (<.001)	.161
NTR	-.06 (.21)	-.06 (.32)	.25 (<.001)	-.01 (.88)	.065
RADAR-Y	.10 (.149)	-.09 (.234)	-.07 (.303)	-.10 (.145)	.032

Educational Achievement

Table S18. *Parental Age Predicting Educational Achievement from Exploratory Results*

Cohort	Age F. β (*p*-value)	Age^2 F. β (*p*-value)	r^2	Age M. β (*p*-value)	Age^2 M. β (*p*-value)	r^2
Gen-R	.15 (<.001)	-.16 (<.001)	.033	.16 (<.001)	-.11 (.001)	.043
NTR	.10 (<.001)	-.01 (.37)	.009	.11 (<.001)	-.02 (.20)	.013

Table S19. *Age Father and Gender Child Predicting Educational Achievement from Exploratory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.15 (<.001)	-.17 (<.001)	.06 (.04)	.036
NTR	.09 (<.001)	-.01 (.45)	-.05 (<.001)	.011

Table S20. *Age Mother and Covariates Predicting IQ from Exploratory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	SES *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.04 (.19)	-.04 (.16)	.47 (<.001)	.05 (.03)	.233
NTR	.05 (.001)	.01 (.28)	.29 (<.001)	-.05 (<.001)	.091

Table S21. *Age Mother and Gender Child Predicting Educational Achievement from Exploratory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.17 (<.001)	-.11 (.001)	.06 (.02)	.047
NTR	.11 (<.001)	-.02 (.22)	-.05 (<.001)	.016

Table S22. *Age Mother and Covariates Predicting Educational Achievement from Exploratory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	SES *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.01 (.66)	-.02 (.68)	.47 (<.001)	.05 (.03)	.232
NTR	.05 (.001)	-.00 (.89)	.28 (<.001)	-.05 (<.001)	.090

Table S23. *Hypotheses for Attention Problems per Rater before and after adjustment for the Covariates Gender and SES.*

	Age Father					Age Mother				
	H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
	$\beta_1 = 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 > 0$	$\beta_1 = 0$ $\beta_2 > 0$	β_1 β_2	$\beta_1 = 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 > 0$	$\beta_1 = 0$ $\beta_2 > 0$	β_1 β_2
Before Adjusting										
Att. Problems										
Child	x				x	x				x
Mother	x		x	x	x	x	x	x		x
Father	x	x	x		x	x	x	x		x
Teacher	x	x			x	x	x	x		x
After Adjusting										
Att. Problems										
Child	x				x	x				x
Mother	x	x	x	x	x	x	x	x	x	x
Father	x	x			x	x	x			x
Teacher	x				x	x			x	x

Note. These hypotheses are based on the results of the exploratory analyses.

Table S24. *Hypotheses for IQ and EA before and after adjustment for the Covariates Gender and SES.*

	Age Father					Age Mother				
	H_1	H_2	H_3	H_4	H_a	H_1	H_2	H_3	H_4	H_a
	$\beta_1 = 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 > 0$	$\beta_1 = 0$ $\beta_2 > 0$	β_1 β_2	$\beta_1 = 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 = 0$	$\beta_1 < 0$ $\beta_2 > 0$	$\beta_1 = 0$ $\beta_2 > 0$	β_1 β_2
Before Adjusting										
IQ	x	x	x	x	x	x	x	x	x	x
EA	x	x	x		x	x	x	x		x
After Adjusting										
IQ	x	x	x	x	x	x	x			x
EA	x	x			x	x	x			x

Note. These hypotheses are based on the results of the exploratory analyses.

Tables for Confirmatory Data

Attention Problems

Table S25. *Parental Age Predicting Attention Problems from Confirmatory Results*

Rater	Cohort	Age F. β (*p*-value)	Age² F. β (*p*-value)	r^2	Age M. β (*p*-value)	Age² M. β (*p*-value)	r^2
Child	**Gen-R**	-.02 (.01)	.00 (.05)	.002	-.02 (.002)	.00 (.40)	.002
	TRAILS	-.04 (.68)	-.02 (.85)	.001	-.01 (.90)	-.02 (.81)	.0001
Mother	**Gen-R**	-.04 (<.001)	.00 (.001)	.006	-.04 (<.001)	.00 (.01)	.007
	NTR	-.04 (<.001)	.00 (.02)	.003	-.05 (<.001)	.00 (.01)	.004
	TRAILS	-.29 (.01)	.10 (.42)	.008	-.42 (<.001)	.15 (.17)	.017
	RADAR-Y	-.20 (.079)	.01 (.276)	.018	-.32 (<.001)	.03 (.13)	.040
Father	**Gen-R**	-.05 (<.001)	.00 (.04)	.01	-.04 (<.001)	.00 (.02)	.005
	NTR	-.03 (<.001)	.00 (.03)	.002	-.04 (<.001)	.00 (.04)	.003
Teacher	**NTR**	-.04 (.06)	.00 (.15)	.000	-.04 (.04)	.01 (.11)	.001
	TRAILS	-.08 (<.001)	.02 (.27)	.018	-.09 (<.001)	.04 (.03)	.031

Note. F. = Father. M. = Mother.

Table S26. *Age Father and Covariates Predicting Attention Problems from Confirmatory Results*

Rater	Cohort	Age β (*p*-value)	Age² β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Child	**Gen-R**	-.01 (07)	.00 (.13)	-.10 (.01)	-.44 (<.001)	.011
	TRAILS	-.05 (.58)	-.01 (.91)	.07 (.56)	.06 (.72)	.001
Mother	**Gen-R**	-.02 (.01)	.00 (.08)	-.38 (<.001)	-.97 (<.001)	.04
	NTR	-.03 (.001)	.00 (.12)	-.36 (<.001)	-1.07 (<.001)	.046
	TRAILS	-.17 (.18)	.09 (.45)	-.66 (<.001)	1.22 (<.001)	.058
	RADAR-Y	-.18 (.10)	.01 (.47)	4.40 (.018)	-1.72 (.118)	.049
Father	**Gen-R**	-.04 (<.001)	.00 (.16)	-.20 (<.001)	-1.10 (<.001)	.042
	NTR	-.02 (.003)	.00 (.16)	-.35 (<.001)	-1.00 (<.001)	.043
Teacher	**NTR**	-.03 (.18)	.00 (.41)	-.59 (<.001)	-4.51 (<.001)	.086
	TRAILS	-.04 (.04)	.02 (.36)	-.17 (<.001)	.24 (<.001)	.108

Table S27. *Age Mother and Covariates Predicting Attention Problems from Confirmatory Results*

Rater	Cohort	**Age** ***β* (*p*-value)**	**Age²** ***β* (*p*-value)**	**SES** ***β* (*p*-value)**	**Gender Child** ***β* (*p*-value)**	r^2
Child	**Gen-R**	-.02 (.04)	.00 (.66)	-.09 (.04)	-.44 (<.001)	.011
	TRAILS	-.03 (.77)	-.01 (.87)	.06 (.61)	.06 (.73)	.001
Mother	**Gen-R**	-.01 (.14)	.00 (.16)	-.37 (<.001)	-.97 (<.001)	.039
	NTR	-.04 (<.001)	.00 (.06)	-.344 (<.001)	-1.08 (<.001)	.047
	TRAILS	-.27 (.021)	.17 (.12)	-.58 (<.001)	1.24 (<.001)	.063
	RADAR-Y	-.30 (.01)	.02 (.252)	3.90 (.034)	--1.72 (.12)	.067
Father	**Gen-R**	-.02 (.03)	.00 (10)	-.21 (<.001)	-1.10 (<.001)	.037
	NTR	-.03 (.001)	.00 (.18)	-.34 (<.001)	-1.00 (<.001)	.043
Teacher	**NTR**	-.03 (.12)	.00 (.27)	-.58 (<.001)	-4.51 (<.001)	.086
	TRAILS	-.05 (.008)	.04 (.017)	-.16 (<.001)	.24 (<.001)	.114

IQ

Table S28. *Parental Age Predicting IQ from Confirmatory Results*

Cohort	**Age F.** ***β* (*p*-value)**	**Age² F.** ***β* (*p*-value)**	r^2	**Age M.** ***β* (*p*-value)**	**Age² M.** ***β* (*p*-value)**	r^2
Gen-R	.36 (<.001)	-.02 (<.001)	.023	.53 (<.001)	-.02 (.002)	.04
TRAILS	2.38 (<.001)	-.79 (.114)	.025	3.25 (<.001)	-.83 (.07)	.049
NTR	.31 (.06)	-.04 (.18)	.011	.37 (.05)	-.04 (.20)	.013
RADAR-Y	.21 (.161)	-.04 (.043)	.023	.29 (.086)	-.07 (.01)	.044

Table S29. *Age Father and Covariates Predicting IQ from Confirmatory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	SES *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.16 (<.001)	-.00 (.16)	5.25 (<.001)	.09 (.82)	.136
TRAILS	.95 (.06)	-.28 (.56)	6.874 (<.001)	-.91 (.29)	.122
NTR	.16 (.36)	-.02 (.58)	3.923 (<.001)	-1.71 (.17)	.078
RADAR-Y	.19 (.213)	-.03 (.079)	-3.78 (.154)	1.078 (.495)	.034

Table S30. *Age Mother and Covariates Predicting IQ from Confirmatory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	SES *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.19 (<.001)	-.00 (.68)	5.07 (<.001)	.12 (.77)	.136
TRAILS	1.64 (<.001)	-.47 (.28)	6.41 (<.001)	.90 (.29)	.129
NTR	.19 (.32)	-.02 (.70)	3.91 (<.001)	-1.68 (.18)	.078
RADAR-Y	.27 (.115)	-.06 (.018)	-3.25 (.216)	1.13 (.471)	.053

Educational Achievement

Table S31. *Parental Age Predicting Educational Achievement from Confirmatory Results*

Cohort	Age F. *β* (*p*-value)	Age^2 F. *β* (*p*-value)	r^2	Age M. *β* (*p*-value)	Age^2 M. *β* (*p*-value)	r^2
Gen-R	.20 (<.001)	-.02 (<.001)	.022	.28 (<.001)	-.03 (.002)	.033
NTR	.17 (<.001)	-.01 (.003)	.007	.23 (<.001)	-.01 (.02)	.011

Table S32. *Age Father and Covariates Predicting Educational Achievement from Confirmatory Results*

Cohort	Age *β* (*p*-value)	Age^2 *β* (*p*-value)	SES *β* (*p*-value)	Gender Child *β* (*p*-value)	r^2
Gen-R	.02 (.70)	-.00 (.74)	4.83 (.53)	.29 (.53)	.000
NTR	.08 (.003)	-.00 (.19)	2.41 (<.001)	-.97 (<.001)	.076

Table S33. *Age Mother and Covariates Predicting Educational Achievement from Confirmatory Results*

Cohort	Age β (*p*-value)	Age^2 β (*p*-value)	SES β (*p*-value)	Gender Child β (*p*-value)	r^2
Gen-R	-.06 (.28)	-.00 (.71)	4.97 (<.001)	.25 (.58)	.214
NTR	.11 (<.001)	-.01 (.15)	2.37 (<.001)	-.96 (<.001)	.077

Bayesian Hypothesis Evaluation and Evidence Synthesis in Confirmatory Data

Attention Problems Regressed on Paternal Age

Both Gen-R and TRAILS strongly supported H_1, the absence of a relation between paternal age and child-reported attention problems. This was found both before and after adding child gender and SES as covariates.

For mother-rated attention problems, Gen-R, TRAILS, and Radar-Y best supported H_1. That is, the absence of a relation between paternal age and mother-rated attention problems. NTR, on the other hand, preferred H_3: a negative linear relation with age with a positive quadratic factor. The NTR plot of this relation showed a wide inverted-U curve, implying that younger and older fathers had children with more attention problems. The Bayesian evidence synthesis showed that over all cohorts, H_1 received most support. Adding covariates did not change the preferred hypotheses. Over all cohorts, age was not associated with mother-reported attention problems when taking the covariates into account.

For father-rated attention problems both Gen-R and NTR best supported H_2, the negative linear relation. After including covariates, Gen-R still supported H_2 best, but NTR now supported the absence of an effect, i.e., H_1. Overall, the negative linear hypothesis was still best supported.

For teacher-rated attention problems, NTR supported the absence of an effect, H1, whereas TRAILS best supported H_2, the negative linear relation. Overall, the negative linear hypothesis was best supported. After correction for covariates, the null hypothesis was preferred by both cohorts.

Attention Problems Regressed on Maternal Age

Both Gen-R and TRAILS strongly supported H_1 for the relation between maternal age and child-reported attention problems. That is, there is no relation between maternal age and child-reported attention problems before and after adding the covariates.

Without covariates in the model, the preferred hypothesis by all cohorts for mother-reported attention problems was H_2, a negative linear relation with age. After including covariates, Gen-R preferred H_1 best, that is, no relation between maternal age and mother reported problems. NTR and Radar-Y still found most support for H_2, the negative linear relation with mother-reported attention problems, whereas TRAILS found substantial support for both H_1 and H_2. Evidence synthesis over the cohorts resulted in most support for H_2, a negative linear relation between maternal age and mother-reported attention problems when taking the covariates into account.

For father-rated attention problems without covariates in the model, Gen-R preferred H_1, whereas NTR preferred H_2, the negative linear hypothesis. Overall, the negative linear relation between maternal age and father-reported attention problems was best supported. After including covariates, Gen-R still preferred H_1, and NTR still H_2, but the best supported hypothesis by both cohorts simultaneously then became H_1: there is no relation between maternal age and father-reported attention problems when taking into account the covariates.

For teacher-reported attention problems, NTR preferred H_1, while TRAILS supported a negative linear relation either with or without a quadratic trend (PMP = .47 and .41 for H_3 and H2 respectively). Over both cohorts, the best supported hypothesis was H_2, a negative linear relation with maternal age. With covariates in the model, TRAILS was ambiguous about the best hypothesis, but NTR preferred the null hypothesis. Hence, over cohorts the null hypothesis was best supported.

IQ Regressed on Paternal Age

Findings across cohorts were mixed. Gen-R found most support for H_3, a positive linear relation with a negative quadratic factor, NTR and Radar-Y found most support for H_1, and TRAILS best supported H_2, a positive linear relation. Over the cohorts, H_3, a positive linear relation with a quadratic effect was best supported. In Gen-R, NTR, and RADAR-Y the form of the quadratic relation was that of a wide inverted U, with slightly lower scores for younger and older fathers (see Figure 2a-2c). TRAILS showed a positive relation that attenuated with older age (see Figure 2d). After including child gender and SES in the model, NTR, TRAILS and Radar-Y best supported H_1, whereas Gen-R found most support for H_2, a positive linear effect. Over cohorts, H_1 was preferred: there is no relation between paternal age and IQ after the covariates in the model.

IQ Regressed on Maternal Age

For maternal age, findings were also mixed over cohorts. Gen-R found most support for H_3, a positive linear effect with a negative quadratic factor. RADAR-Y preferred H_4, a negative quadratic factor but no linear effect. The Gen-R plot showed a positive effect of age that attenuated over time, whereas the RADAR-Y plot showed an inverse-U curve with younger and older parents having children with lower IQ scores. NTR preferred H_1, and TRAILS found most support for H_2, a positive linear relation. Over the cohorts, H_2 and H_3 were best supported (PMP H_2 = .51, PMP H_3 = .48). After including covariates in the model, Gen-R and TRAILS found most support for H_2, a linear relation with maternal age, whereas NTR found most support for H_1, and Radar-Y supported both H_1 (PMP = .38) and H_2 (PMP = .34). Overall, H_2, the positive linear relation between maternal age and IQ was best supported by all cohorts.

Educational Achievement Regressed on Paternal Age

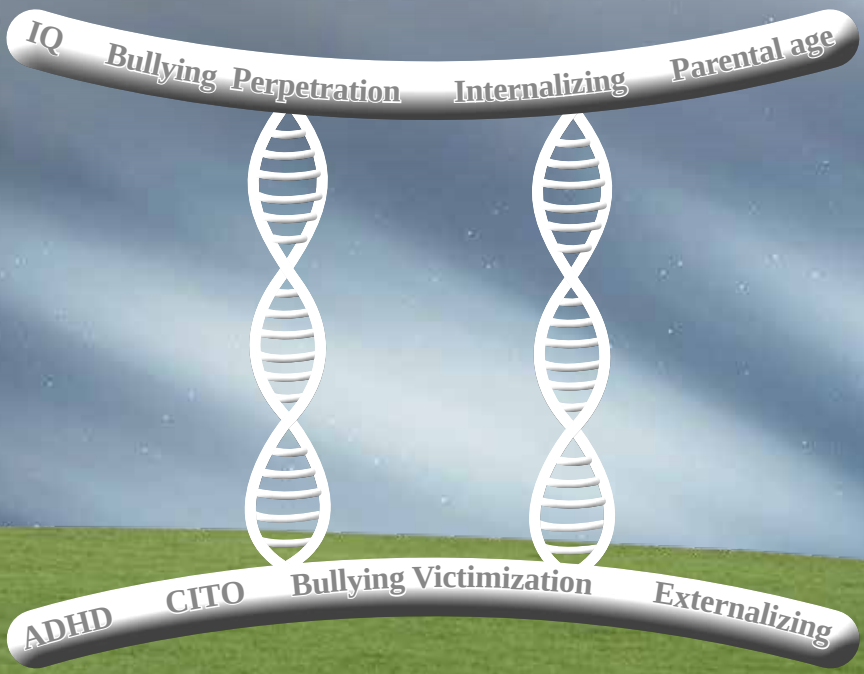
Both Gen-R and NTR best supported H_3, a positive linear relation with a negative quadratic factor between paternal age and EA. Plots of this relation (see Figure 3a-3b) showed a wide inverted-U curve, implying that younger and older fathers had children with lower EA scores. After including child gender and SES in the model, Gen-R preferred H_1, whereas NTR found support for both H_1, no relation between paternal age and EA, and H_2, a positive linear relation between paternal age and EA. Over cohorts, the null hypothesis of no relation between paternal age and EA was best supported when including the covariates in the model.

Educational Achievement Regressed on Maternal Age

For maternal age, Gen-R found most support for H_3, a positive linear relation with a negative quadratic trend. The Gen-R plot showed a positive effect of age that attenuated over time (see Figure 3c). NTR, on the other hand, found most support for H_2, a positive linear relation. Overall, most support was found for H_3 (see Figure 3d for the NTR plot). After correction for covariates, Gen-R preferred H_1, whereas NTR preferred the positive linear relation. Overall, most support was found for H1, there is no support for a relation between maternal age and EA after including child gender and SES in the model.

Chapter 6

Summary and General Discussion



Summary

This dissertation focused on 1) bullying in primary school, and 2) the influences of parental age on childhood cognitive development and behavioral and emotional problems. For both domains, I looked at the role of protective and risk factors. With respect to bullying, data of twins were analyzed in two ways. First, in twin data we studied the effect of close companionship on bullying. Second, the classical twin design made it possible to advance knowledge about the etiology of differences between children by estimating the contribution of genetic and environmental factors to bullying. Regarding parental age, the aim was to advance knowledge about the influence of advanced parental age on offspring's externalizing- and internalizing problem behavior, attention problems, and cognitive functioning. For this aim I analyzed data from a large number of children through collaborations with other multiple childhood cohorts in the Netherlands. In this chapter I will first summarize the main findings of the two studies regarding bullying behavior and next of the two studies on the influence of parental age.

1. Bullying

The first part of my thesis was about bullying behavior in twins during primary school and addressed two issues: 1) the influences of possible risk factors on the prevalence, and 2) the causes of familial resemblance.

Chapter 2 was about risk factors regarding bullying behavior in twins. Based on previous research, a lot of questions remained regarding bullying in twins. The factors that were investigated in this chapter can be divided into twin specific and non-twin specific factors. For twin specific factors, I investigated whether the risk for bullying perpetration and bullying victimization differed for monozygotic- versus dizygotic twins, same-sex versus opposite-sex twins, and twins attending the same versus separate classrooms. In follow-up analyses, I also investigated two possible interaction effects: 1) whether an effect of classroom sharing differed for monozygotic- and dizygotic twins, and 2) whether an effect of being a same-sex versus an opposite-sex twin pair changed as children age. Regarding non-twin specific factors, I investigated whether the prevalence of perpetration and victimization change as children age and whether the prevalence rate of perpetration and victimization in boys differed from that in girls. In addition, an important question was about whether having a co-twin protected twin-children from bullying or being bullied. This question whether or not close companionship is protective was investigated by comparing the prevalence of perpetration and victimization in twins with the prevalence of their non-twin siblings. By using twins and singleton siblings from the same families, both groups match each other on important family background variables.

All research questions were investigated for both bullying perpetration and bullying victimization in a sample of ~ 8,000 twin children and ~1,400 singleton children. Bullying was rated by their teachers at ages 7, 9-10, and 12. Teachers answered four items about perpetration and four items about victimization. The items for victimization assessed (1) 'how often has the child been victimized in the past couple of months? (in general)', (2) 'how often has the child been teased, laughed at, or called names in the past couple of months? (verbal victimization)', (3) 'how often has the child been physically victimized, such as being hit, kicked, and pushed in the past couple of months? (physical victimization)', and (4) 'how often has the child been excluded by other children, ignored, or have other students spread false rumors? (relational victimization)'. Each item was scored on a five-point scale, from *never, once or twice, two or three times a month, about once a week, and several times a week.* Perpetration was assessed with the same items, but in the active form.

My study showed that close companionship was not protective, based on the finding that twin children are as much involved in bullying as their non-twin siblings. For both twins and singletons, teachers reported that 36% of the children bullied their peers moderately to severely in the last couple of months (i.e., at least once), and 35% of the children suffered moderately to severely from victimization. The twin-specific factors revealed that being fraternal or identical twins, or being part of a same-sex or opposite-sex twin pair does not affect the prevalence rates. The most important twin specific finding, however, is that classroom sharing appeared to be a protective factor regarding victimization. A subsequent analysis showed that this finding was restricted to girl-girl twins. In other words, female twin pairs placed together in the same classroom do, on average, not bully more often, but are less often victimized by others. Based on this finding, we conclude that assigning female twins to the same classroom may act protectively. More general, for girls, this suggests a protective effect of having a close companionship in the same classroom. The non-twin specific factors showed that children around age 10 are at highest risk to be involved in bullying and that boys are more often involved in bullying, either as bully or victim.

After taking into account the general effects on bullying, large individual differences remain. In **Chapter 3**, I addressed the question to what extent these individual differences in the risk of bullying are caused by genetic- and environmental factors. Here I considered these factors as latent concepts containing all genetic and environmental variation between children and estimated the extent to which these factors influence why some children are involved in bullying and others not. This question could be addressed by making use of the classical twin design, that includes mono- and dizygotic twin pairs.

I estimated the relative contribution of genetic and environmental factors on different forms of bullying perpetration, bullying victimization, and their association. Teachers rated ~8,000 twins on their general, physical, verbal and relational bullying behavior by using the same items as in chapter 2.

The teachers reported that 34% of these children were involved as a bully, victim, or both. The heritability of perpetration was ~70%, for victimization the heritability was ~65%, similar for boys and girls, yet both were somewhat lower for the relational form. More specifically, for both boys and girls the heritability estimate of general perpetration was 72%, for verbal perpetration the heritability was 73%, for physical perpetration 71%, and for relational perpetration 68%. For victimization these estimates were respectively 62%, 64%, 70%, and 55%. Shared environmental influences for perpetration and victimization were modest (ranged from 2%-18%) and were more pronounced among girls. Bullying perpetration and bullying victimization were highly correlated.
The correlations in our sample ranged from .59 (for the relational form) to .85 (for the physical form).

The association between being a bully and being a victim was mostly explained by shared genetic factors for the general (~65%), verbal (~71%) and physical (~77%) forms and mostly by environmental factors for the relational form (~60%). This translates into genetic correlations of .50 for general bullying, .62 for verbal bullying, .86 for physical bullying, and .26 for relational bullying.

2. Parental Age

The second part of my thesis was about the influences of parental age on two important aspects of child development: 1) externalizing and internalizing problem behavior, and 2) attention problems and cognitive functioning.

In the literature, the effects of advanced parenthood on neurodevelopmental disorders, like autism and schizophrenia, are well established (Merikangas, 2016; 2017). That is, offspring of older parents are more at risk to develop these disorders. However, for other child characteristics less is known about the effect of advanced parenthood. **Chapter 4** reported on the influence of parental age on offspring internalizing and externalizing problems. Based on previous studies regarding neurodevelopmental disorders, our expectation was that the adverse effects of older parents might extend to offspring problem behavior. We analyzed the influence of advanced mother- and fatherhood within four large Dutch population-based cohorts, with a total sample of ~33,000 10-12 aged children.

The cohorts that contributed to this study were the Netherlands Twin Register (NTR, see van Beijsterveldt et al., 2013; from all regions in the Netherlands), Generation R (Gen-R, see Kooijman et al., 2016; city of Rotterdam in the Netherlands), the Research on Adolescent Development and Relationships-Young cohort (RADAR-Y, see Crocetti et al., 2017; province of Utrecht and four large cities in the mid-west of the Netherlands), and the Tracking Adolescents' Individual Lives Survey (TRAILS, see Oldehinkel et al., 2015; the Northern regions of the Netherlands). Externalizing and internalizing problems were rated by multiple informants: mothers, fathers, teachers and the children themselves. Each cohort had data available for at least one informant. Both outcomes were assessed with the ASEBA questionnaires, which include standardized instruments for child self-reports, parent reports, and teacher reports. We executed cross-validation analyses by using the first random half of the data for generating hypotheses and by using the other half of the data for testing these hypotheses. Cutting the whole dataset of each cohort into two independent datasets avoids "double dipping". That is, in this way the informative hypotheses are not generated and evaluated by the same dataset. The exploratory results of the various cohorts showed that it might be possible that 1) age had a negative linear effect and no quadratic effect, or 2) that age has a negative linear effect with a positive quadratic effect. Hypotheses representing "no effect" and "all other effects than specified in the informative hypotheses" were also tested in the confirmatory phase. Each cohort evaluated this same set of hypotheses. Bayesian evidence synthesis was applied to summarize the results of the multiple cohorts.

Based on the confirmatory results, we can state that there was evidence of a robust (i.e., "over cohorts") negative linear relation between parental age and externalizing problems when the analyses were based on parent reports, indicating that children from older parents show less externalizing problems. In teacher-reports, this relation was largely explained by socio-economic status. Child-reported data showed no effect with parental age. Parental age had limited to no association with internalizing problems. These results indicate that there is no harmful effect of advanced parenthood on offspring's externalizing and internalizing problem behavior. For externalizing problem behavior, there even is a beneficial effect, both before and after including SES.

The method we applied in this study thus contained four steps: 1) creating exploratory and confirmatory datasets, 2) generating informative hypotheses using the exploratory dataset, 3) evaluating these informative hypotheses using Bayesian hypothesis evaluation, and 4) using Bayesian evidence synthesis to summarize all results of the multiple cohorts into an overall "robust" result.

In traditional null hypothesis significant testing it is not possible to quantify the support for the null-hypothesis, which appeared an important hypothesis in our study. In our study the generated informative hypotheses are evaluated to this traditional null-hypothesis and the alternative hypothesis. Consequently, this Bayesian method should increase the credibility of our results. Since our method is based on quantifying support for each informative hypothesis instead of rejected or not-rejecting the null-hypotheses, it should also reduce publication bias. Classical meta-analyses are biased since a lot of studies with null-findings are not published. Another important strength of Bayesian evidence synthesis over classical meta-analyses is that our method enabled us to combine the results of the multiple cohorts into robust overall results, even when the multiple cohorts used different measurement instruments for the same concepts.

Chapter 5 investigated the effects of parental age on neurodevelopmental outcomes that are more common than autism and schizophrenia, like attention problems and cognitive functioning. Here, child-, father-, mother- and teacher-rated attention problems (N ~ 38,000), intelligence (N ~ 10,000) and educational achievement (N ~ 17,500) were analyzed for children from NTR, Gen-R, TRAILS, and RADAR-Y. Data for attention problems and intelligence (IQ) were available for each cohort. Standardized educational achievement data, measured by the "CITO-test" (Citogroep, 2019), were available for two cohorts. The "CITO-test" is a 3-day nation-wide standardized test for children at the end of primary school (around age 12). Around 75 per cent of schools in the Netherlands took part. For attention problems, each cohort had data available for at least one informant. The method we applied to analyze these data was the same as applied in chapter 4. That is, first informative hypotheses were generated based on the exploratory part of the data, after which these hypotheses were evaluated based on the confirmatory part of the data. The random first half of the data discovered that age 1) might have a negative linear relation and no quadratic relation, 2) might have a negative linear relation and a positive quadratic relation, or 3) might have a positive quadratic relation, but no linear relation. The hypotheses for cognitive functioning were the reverse. These informative hypotheses were tested in the confirmatory phase. Based on Bayesian evidence synthesis, the confirmatory analyses showed that older parents have offspring with fewer attention problems and younger parents have offspring with more attention problems. For IQ and educational achievement, the age of the mother also showed a positively and linearly effect. For fathers, however, their age had an attenuating positive relation with educational achievement, and an inverted U-shaped relation with IQ. This inverted U-shaped effect means that younger and older fathers are disadvantaged.

We thus conclude that, in general, there were hardly any disadvantages for offspring of older parents with respect to the neurodevelopmental conditions attention problems, IQ, and educational achievement. We even showed that advanced parental age is mostly advantageous for attention problems and educational achievement. These associations mostly disappeared after including SES, indicating that SES had an important role in the relation between parental age and offspring neurodevelopmental outcomes.

General Discussion

As an overall aim of my dissertation I wondered which influences make some children vulnerable to face developmental difficulties and make others resilient. The aim of this thesis was to expand knowledge about two issues regarding child development: 1) bullying during primary school, and 2) influences of parental age on child development.

1. Bullying

In this part of my thesis, I investigated risk factors for bullying in **Chapter 2** and genetic and environmental influences on bullying in **Chapter 3**.

1.1 Risk Factors

The body of literature regarding bullying mostly addressed the effects of general factors, ranging from individual (e.g., age and gender) to contextual (e.g., parenting). Not many studies were done to investigate twin specific risk factors or to compare twins with singletons, even though twin children constitute 1/40 of all children. I will highlight and discuss the two most interesting and remarkable findings regarding risk factors for bullying in twins, which are about twin-singleton differences and the protective effect of classroom-sharing.

Twin-singleton differences

The first key finding of **Chapter 2** is that twin children are as much involved in bullying as their singleton, i.e., non-twin, siblings. Previous studies that tried to answer this "twin-singleton" question showed mixed results *(singletons at higher risk:* Barnes & Boutwell, 2013; no effect: Oshima et al., 2010; *twins at higher risk:* Weissenberg et al., 2007). These studies, however, were all based on unrelated singletons. Important related family factors in these unrelated singletons could have differed from that in the twin group, which hampers the twin-singletons comparisons of previous studies. This means that previous studies were thus not able to distinguish between real effects or effects caused by differences in important background characteristics of the twins and singletons.

The strength of our study is that we collected, via teachers, data on the brothers and sisters of twins, instead of on unrelated singleton children from different families. That is, we employed a within family design that included twins and their singleton siblings. Within-family analyses are based on a research design that avoids confounding. That is, by comparing twins to their singleton sisters and brothers, we made sure that we controlled for influences of maternal and family characteristics (e.g., the socio-economic environment) that might be related to bullying. For example, Jansen et al. (2012) showed that the prevalence of bullying is influenced by family SES. That is, children with low family SES have an increased risk of being a bully or bully-victim. So, we made sure that both groups match each other on important family background variables. This method of matching important family background information of twins and singletons enabled us to detect true effects. Consequently, we can conclude that there are no differences regarding the bullying prevalence rates of twins and singletons. The implication of this finding is that it is likely that our conclusions based on twin research (like age and gender effects) are therefore generalizable to the population at large.

Classroom Effect

One of the most notable finding of **Chapter 2** is that female twin pairs who attend the same classroom do not bully less or more than separated twins, but they are less victimized. The only previous study that investigated this research question was the study of Lamarche and colleagues (2006). Their result pointed towards a protective effect of sharing a classroom, but their study was possibly too small to reach significance for this finding. Our result, based on a much larger group of twin pairs, is in agreement with the result of Lamarche et al. (2006).

Previous studies found very few harmful or beneficial effects of classroom sharing on other traits, like school performance or problem behavior (van Leeuwen, van den Berg, van Beijsterveldt & Boomsma, 2005). There are also studies showing that separated twins had more internalizing problems and lower reading scores, suggesting that classroom sharing might be beneficial, but we have to take in mind that all these effects were weak (e.g., Lamb, Middeldorp, van Beijsterveldt & Boomsma, 2012; Tully, Moffit, Caspi, Taylor, Kiernan & Andreaou, 2004). In agreement with these results, we found that girl-girl twins attending the same classroom are less often victim of bullying than those in separate classrooms, suggesting that for victimization being together in the same classroom may also act protectively.

Together, these findings imply that the placement of twins in the same classroom might be beneficial regarding victimization and is not harmful for other important behavioral and cognitive outcomes. This is very relevant and important to twin families, primary schools, and educational policies, since classroom placement is a malleable factor and teachers and parents can thus have an influence. Although school may have a set policy regarding classroom placement of twins, this protective effect of classroom sharing for girl-girl twins should be taken into account when twins enter primary school.

1.2 Etiology

Only very few studies investigated the causes of individual differences in bullying within a genetics context. Most of them focused on victimization, while perpetration and their association was only investigated once (by Ball et al., 2008). Although we know from earlier studies that bullying behavior comes in different forms, ranging from kicking (physical) and name-calling (verbal) to spreading rumors (relational), all genetically-sensitive studies so far focused on bullying behavior in general. The one exception is the study of Eastman et al. (2018), who investigated the heritability of various forms of victimization. Our data enabled us to investigate the genetic and environmental influences on different types of bullying perpetration, victimization, and their co-occurrence (characterizing bully-victims).

Etiology of the different Forms of Perpetration and Victimization

We found that all forms of bullying perpetration and victimization were substantially heritable. Eastman et al. (2018) investigated self-reported verbal, physical, relational, and property victimization and showed that these heritability estimates ranged from 23% (for attacks on property) to 42% (for physical victimization). Their finding that physical victimization is most heritable is in agreement with our findings. Their study, however, suffered from limited power (N=306 pairs) and therefore they could not investigate whether the heritability estimates differed for boys and girls. They also did not investigate different forms of perpetration.

Our large sample (N ~ 4,500 twin pairs) and data on both perpetration and victimization enabled us to investigate both remaining research questions. We showed that there were no differences in heritability estimates for boys and girls. Regarding perpetration, we cannot compare our results of the various forms with previous studies, since there are no previous studies. However, as we compare it with the general perpetration heritability Ball et al. (2008) showed, then we see that they are comparable (~70% in our study, compared to 61% in their study).

Etiology of their Associations

The co-occurrence of perpetration and victimization was mostly due to genetic factors for general, verbal, and physical bullying, but mostly due to environmental factors for relational bullying. The only previous study (Ball et al., 2008) showed a (phenotypic) correlation of .25 between general bullying perpetration and general bullying victimization and they went on to show that this correlation was only due to genetic factors. Although in our study, the correlation for the general item between perpetration and victimization was much higher (~.65), their finding that this correlation was only due to genetic factors is to a large extent in agreement with our finding. We found that this correlation was mostly caused by genetic factors. However, we also showed that the causes of the associations differed for the various subtypes (i.e., lower influences of genetic factors for the relational form) and therefore we cannot easily compare our results with previous studies that might have combined information of different subtypes into one single measure. For instance, Ball et al. (2008) assessed the construct general victimization with items including verbal, physical and relational victimization.

1.3 Conclusions Regarding Bullying

Together, the results of **Chapter 2** imply that the factors influencing bullying indeed range from individual (like gender) to contextual (like classroom-sharing). Moreover, the results of **Chapter 3** implicate that substantial genetic influences are responsible for the individual differences regarding bullying behavior (for bullies, victims and bully-victims). Teachers, for example, can confirm that some children are more vulnerable for being victimized, especially children that are different based on their appearance or behavior. We know from twin research that most of the physical appearance and behavior is moderately to highly heritable. In addition, a recent study showed that children with a genetic vulnerability for attention-deficit hyperactivity disorder (ADHD), depression, higher BMI and lower IQ, are at a higher risk for exposure to bullying (Schoeler et al., 2019). This genetic vulnerability might explain the genetic influences on bullying. This genetic vulnerability, however does not mean that bullying is not modifiable. There are evidenced-based interventions that reduce school bullying. Interventions have been shown to be most effective when the whole school is involved (e.g. Menesini & Salmivalli, 2017; Vreeman & Carroll, 2007), an example of such a program is the KiVa Anti-Bullying Programm (Salmivalli, Kaukiainen & Voeten, 2005). Positive teacher-child relationships might reduce the vulnerability of children and the detrimental effects victimization has. Teachers that clearly communicate their antibullying attitude to the children in their classroom might be of influence as well (e.g. Menesini & Salmivalli, 2017). Moreover, the findings of Schoeler et al. (2019) implicate that prevention programs should address preexisting vulnerabilities in order to avoid repeated exposure to bullying.

2. Parental Age

In the second part of my dissertation, I collaborated with multiple large childhood cohorts in the Netherlands. These cohorts work together within the Consortium on Individual Development (CID). CID is an NWO-funded consortium in which different universities and institutes collaborate, combining expertise of multiple disciplines. This consortium aims to understand the factors that influence individual differences, ranging from child characteristics to environmental factors. Our collaboration within this large consortium turned out to be highly valuable. We wrote two papers together which should be of considerable interest to a large audience. These papers are of great value due to the advanced statistical approach and large datasets and can therefore extend existing knowledge about the influence of parental age on childhood development.

With the unique data of this large collaborative effort, I focused on the influences of parental age on offspring's socio-emotional and cognitive development. The rationale for these two studies was that people in western societies are more and more postponing parenthood and this might have consequences for children's (mental) health. In the Netherlands, for example, the mean age at first birth was 24 in 1970, while nowadays this is much higher, around 30 (CBS, 2019). In Figure 1, it can be seen that the children born to mothers aged 35 years or above were in 1950 mostly the fourth (or subsequent) born child, while in 2015 these children were mostly the first or second born child. In other words, this figure shows that in 1950 and 1975 mothers started a family earlier compared to mothers in 2015.

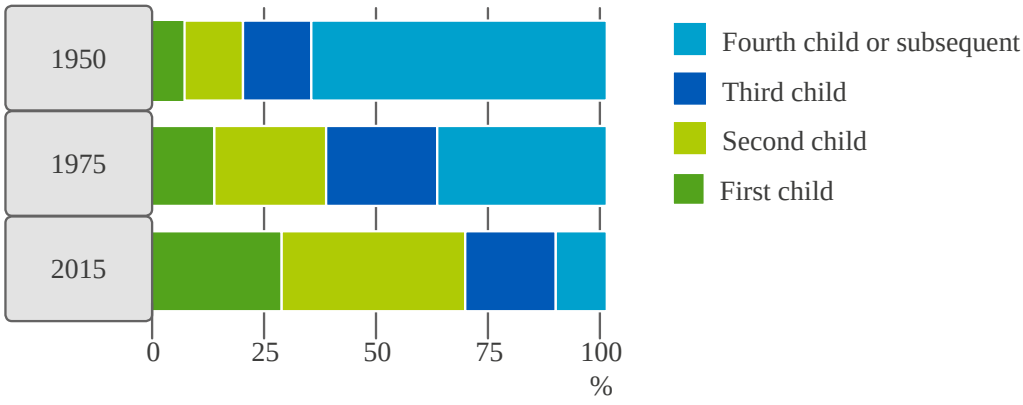


Figure 1. *Live births to mothers of 35 years or above, for three different birth cohorts.*

In line with Figure 1, Figure 2 also shows that in general people tend to postpone parenthood until later in life. The total number of children born fell rapidly after the widespread availability of contraception around 1970. Regarding the trend of delaying parenthood, for example, the yellow band in Figure 2 demonstrates that the number of mothers of 30-35 years of age has considerably increased. That is, from approximately 25 per 1.000 mothers in 1950 to approximately 50 per 1.000 mothers in 2015. In other words, the number of mothers aged 30-35 has doubled between 1950 and 2015. Figure 2 (bottom blue band) also shows that nowadays there are very few teen mothers (<20 years of age) in the Netherlands. Since in the Netherlands teen mothers barely exist, in the following sections I will focus on the findings and implications for older parents.

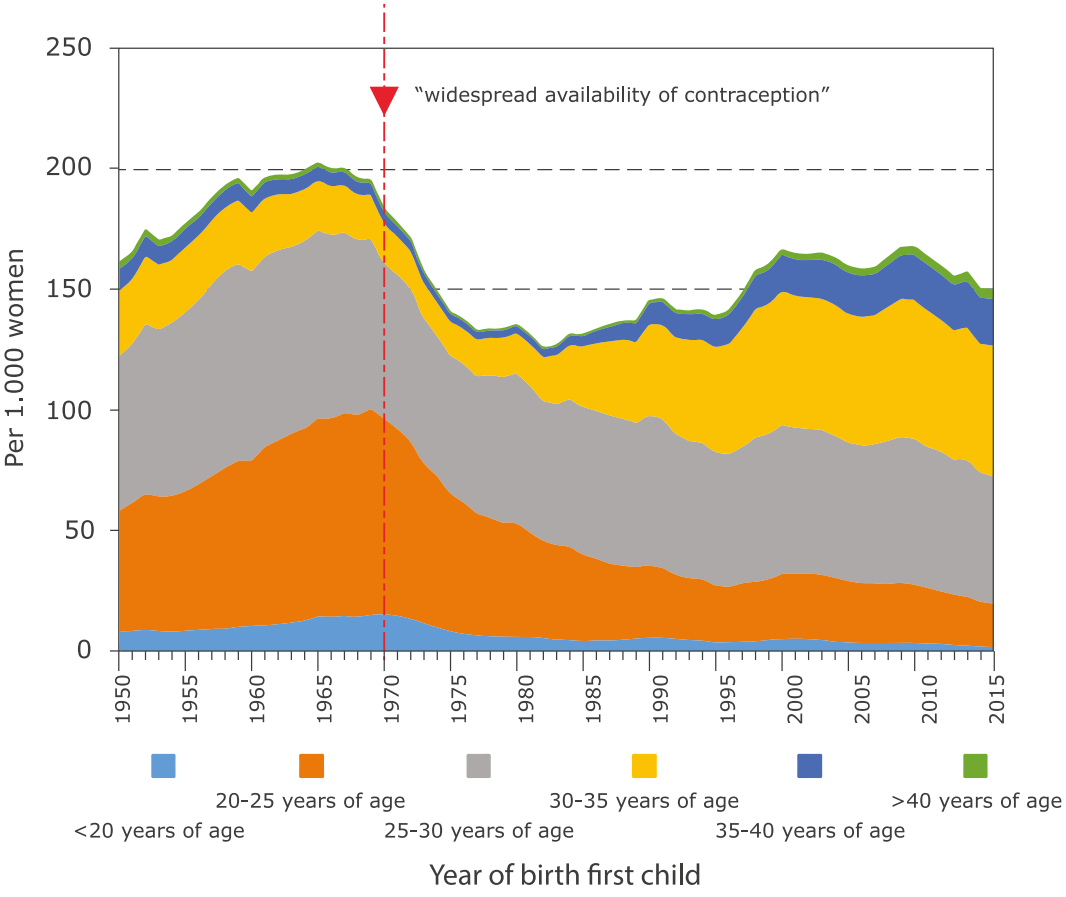


Figure 2. *Number of first born children per 1.000 mothers, across maternal age category and year of child birth.*

Older parents are disadvantaged from a biological point of view, for example due to a higher number of *de novo* mutations (e.g., Malaspina, 2001). It is well established that children of older parents (especially investigated for older mothers) have an increased risk for rare severe neurodevelopmental disorders, like autism, schizophrenia, and Down Syndrome (e.g., Merkiangas, 2016, 2017). Therefore, concerns are growing about the postponement of starting families. In this part of my thesis, I aimed to investigate whether the adverse effects of advanced parenthood for rare severe disorders extent to the full range of socio-emotional and cognitive skills. Therefore, I investigated externalizing- and internalizing problem behavior in **Chapter 4** and attention problems and cognitive functioning in **Chapter 5.**

2.1 Externalizing and Internalizing Problems

In general, we found that offspring of older parents had fewer externalizing problems. It should however be noted that the results were rater dependent. That is, parents and teachers reported fewer externalizing problems for offspring of older parents. Based on the self-reports of the children no effect was found, but this can be due to the limited ability of 10-year-old children to reliably report their behavior. For teacher-reported data, the positive relation mostly disappeared after including information about the socio-economic status (SES) of the parents. This was not the case for parent-reported problem behavior, meaning that the favorable effect of parental age is rater dependent and not solely due to the socio-economic status of the parents.

Only a number of studies have previously investigated the influence of parental age on externalizing and internalizing problems and showed mixed results (for a review, see Tearne, 2015). In addition, it remained unknown whether the effects of advanced maternal and paternal age were the same. Regarding externalizing problem behavior, most studies reported that offspring of older mothers have fewer problems (for review, see Tearne, 2015), which is in line with our result. We showed that this effect extends to older fathers. Regarding internalizing problems, there was even less known. Since there was little comprehensive evidence from previous studies, we aimed to advance knowledge about the effects of advanced parental age on offspring's externalizing and internalizing problem behavior with our advanced method and large cohorts in order to inform worrying (future) parents about possible harmful effects.

Our findings implicate that there were thus no harmful effects of older parents on offspring's externalizing and internalizing problem behavior. Having older parents might even be beneficial regarding externalizing problems.

2.2 Attention Problems and Cognitions

Overall, we showed that offspring of older parents tended to have offspring with fewer attention problems. For cognitive functioning, offspring of older mothers had on average higher IQ and educational achievement scores. Regarding paternal age, offspring of older fathers tended to have higher educational achievement scores, but this effect attenuated with older age (i.e., curvilinear plateau). Offspring of both older and younger fathers had lower IQ scores. Most of these relations, however, disappeared when taking the socio-economic status of the parents into account.

Previous studies regarding children's neurodevelopment showed mixed results. For instance regarding the effects of paternal age on offspring ADHD. Children of older fathers have been found to have a higher risk for ADHD in the study of D'Onofrio et al. (2014), equal risk in the study of Mikkelsen et al. (2016), and reduced risk in the study Chudal et al. (2015). These mixed findings were mirrored in the literature regarding the effects of parental age on cognitive function. All these previous studies used different methods, analytical strategies and/or control variables, which hampers drawing conclusions. Our method, on the other hand, enabled us to combine evidence from four large cohorts and hence obtain robust results, even though cohorts differed in type of measure. This study enabled us to expand the message for worrying (future) parents in Chapter 4 with information about the effects of advanced parental age on aspects of childhood development beyond rare neurodevelopmental disorders.

Our findings implicate that there were no harmful effects of older parents on offspring's attention and educational achievement. Only for IQ possible harmful effects of advanced fatherhood were found. For offspring's attention and educational achievement, the effects of older parenthood might even be beneficial. Taken together, this chapter showed us that advanced parental age is hardly disadvantageous, and mostly advantageous, with respect to attention problems, educational achievement and IQ.

2.3 Conclusions Regarding Parental Age

In the two discussed studies (Chapter 4 and Chapter 5), we reported that offspring of older parents fare better, but importantly, most effects disappeared after correcting for SES of the parents. This suggests that these positive effects are due the interplay between the genetic and environmental effects in SES that are transmitted from parents to their children. SES reflect the family's economic position (i.e., income and education) and is associated with different characteristics of the parents, like the parent's attention, cognitive abilities, and occupational level.

We know that in general most older parents have a higher SES and are highly educated. This implies that the beneficial effect of advanced parenthood seems especially driven by parental educational level.

The influence of SES might be explained in two ways. The first way is that SES might influence parental age, which in turn influences the offspring outcomes, meaning that people of low SES start families earlier which indirectly influences childhood outcomes. For instance regarding ADHD, parents with ADHD have a tendency for lower SES and their more risky behavior might result in early pregnancies. The children of these younger parents have lower family income (SES) and a parental history of ADHD, which both make them on average more likely to develop ADHD compared to other children. Russell, Ford, Williams and Russell (2016) indeed showed that family income and parental history of ADHD are strong predictors of the prevalence of ADHD in the offspring generation. Parental age is thus a mediator between parental SES and offspring outcomes. That is, ADHD in offspring is not due to the age of the parents, but due to the transmitted environmental and genetic liability. Secondly, SES may reflect a more general liability that influences both age at which people start families and offspring outcomes, without causality between parental age and offspring outcomes. This general liability could reflect both the genetic- and environmental transmission from parents to offspring.

Regarding the environment, parents of higher SES obviously have more resources to provide their children with a more stimulating environment. For instance van Bergen et al. (2017) showed that parental level of education, a measure that is highly correlated with SES, and the number of books in the home are moderately correlated (r=.45). So, highly educated parents tend to provide a more stimulating home environment, and transmit more genetic variants to their children that are associated with higher educational attainment. This phenomenon is referred to as "passive" gene-environment correlation; the genes that are transmitted from parents to offspring are not independent from the environments that these same parents provide to their offspring. In fact, the effect of the number of books in the home on children's reading ability mostly disappeared after controlling for genetic influences, leaving little room for genuine environmental effects (van Bergen et al., 2017). This is in line with the finding that parent-child resemblance for reading ability seems to be due to genetic rather than cultural transmission (Swagerman et al., 2017).

The phenomenon of a passive gene-environment correlation may also apply to other traits: for example, parents with ADHD have an increased risk of impulsive behavior and unplanned early pregnancies (Ostergaard, Dalsgaard, Faraone, Munk-Olsen & Laursen, 2017).

These parents transmit their genetic vulnerability for ADHD to their offspring as well as a more chaotic and unstructured environment. More offspring of younger than of older parents may thus have a genetic predisposition to develop ADHD.

Although we do not exactly know which factors cause the beneficial effects of parental age that I found in my PhD research, we can say that it seems that advantages of advanced parenthood might outweigh the biologically disadvantages of advanced parenthood. Based on beneficial effects of parental age, people may think that future parents (especially highly educated people) do not need to worry about the development of their offspring and that it might be better to start a family at a later age. However, one should keep in mind that these results are *general effects* (i.e., in the population at large). I do not want to state that advanced parenthood is always "better", since the biology of ageing still seems to put older parents in an unfavorable position with regard to their offspring's physical and mental health (e.g., Malaspina, 2001). It is well established that offspring of older parents are at higher risk for serious neurodevelopmental disorders, like autism (Merikangas, 2016; 2017) event though the mechanisms that underly the association are still debated. It is also important to mention that the older parents in our (and comparable) studies were indeed parents; some people who postpone parenthood remain involuntary childless (te Velde, Habbema, Leriodon & Eijkemans, 2012), which can have a big psychological impact (Lechner, Bolman & van Dalen, 2006). In Figure 3, taken from the study of te Velde et al. (2012), it can be seen that permanent involuntary childlessness (PIC) in Europe approximately doubled since the 1970's. In the Netherlands, this percentage increased from approximately 2.5% in 1970 to approximately 6.5% in 2007.

My aim was to investigate whether the well-established adverse effects of advanced parenthood for severe neurodevelopmental disorders also would extend to more common neurodevelopmental problems. Based on **Chapter 4** and **Chapter 5,** it is clear that there were hardly any harmful effects of advanced parental age on internalizing and externalizing problem behavior, attention problems, IQ, and educational achievement. Stated differently, offspring of older parents do, on average, equally well or better than offspring of younger parents. The finding that offspring of older parents perform better was mostly found before taking the SES of the parents into account. The finding that there is no effect of parental age was mostly found after taking the SES information into account. Both findings, however, imply that there is no harmful effect for postponing parenthood for the socio-emotional and cognitive development of children and thus parents that started their family at a later age do not have to worry regarding these childhood outcomes.

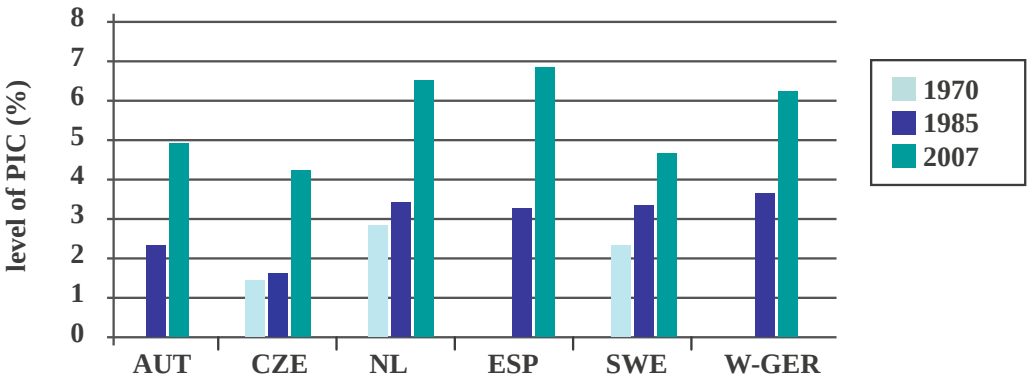


Figure 3. *The effect of postponement of first childbirth on permanent involuntary childlessness (PIC in %) for different European countries. AUT = Austria, CZE = the Czech Republic, NL = the Netherlands, ESP = Spain, SWE = Sweden, and W-GER = West Germany. (This figure is taken from the study of te Velde, Habbema, Leriodon & Eijkemans, 2012).*

The strength of the last two empirical chapters of my thesis is its methodological approach which employed a method to estimate robust effects by using Bayesian synthesis evidence. With the different measures from the four cohort studies, a meta-analysis approach which often requires to have the same measurement method in different studies would have been less optimal. In contrast, the Bayesian method enabled us to estimate robust effects, even though some child characteristics might have been assessed with different measures. Another strength was the large sample sizes of the four different cohorts. Thus, we contribute to the knowledge about parental age effects on childhood development by using a strong methodological approach.

3. General Conclusion

The overall question of my thesis was: which influences make some children vulnerable to face developmental difficulties and make others resilient? Within this overarching question, I aimed to expand knowledge about bullying in primary school and influences of parental age. For bullying, I conclude that there are general risk factors which make some children more vulnerable to be involved, like being a boy. Besides, I also found protective factors, like classroom sharing for girl-girl twins and I take the close bond that twins can have as a model for the protective effects of 'close companionships'. Even after accounting for these general effects, large individual differences remained, which we showed to be mostly caused by genetic differences between children.

In the second part of my thesis, I showed that offspring of older parents tend to have fewer behavioral- and neurodevelopmental problems and higher cognitive functioning. This effect was mostly due to the SES of their parents. In other words, also characteristics of parents, like their age at which they start families and education, influence multiple aspects of their offspring development.

To situate these results within the Consortium for Individual Development (see Figure 4), we can indeed say that children's biological predisposition, children's characteristics and children's rearing environment influence why some children face more difficulties during their development than others.

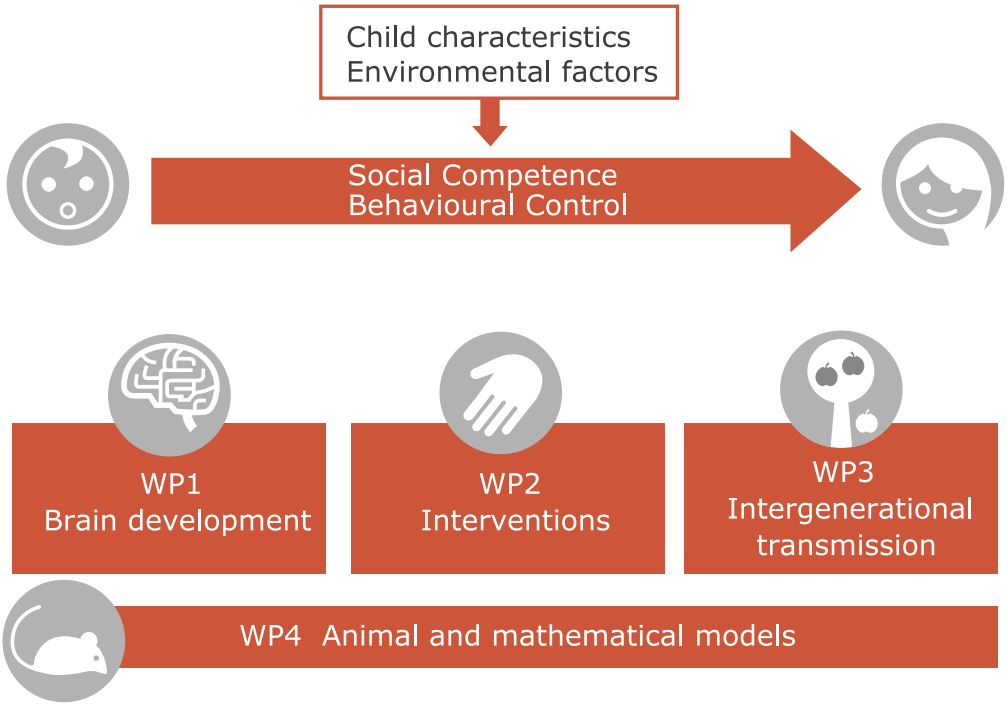


Figure 4. *The overview of the CID consortium.*

References

Ball, H. A., Arseneault, L., Taylor, A., Maughan, B., Caspi, A., & Moffitt, T. E. (2008). Genetic and environmental influences on victims, bullies and bully victims in childhood. *JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY*,49(1), 104-112.

Barnes, J. C., & Boutwell, B. B. (2013). A demonstration of the generalizability of twin-based research on antisocial behavior. *BEHAVIOR GENETICS*, 43(2), 120-131.

Centraal Bureau voor de Statistiek. (2019, January 23). Vrouwen steeds later moeder. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2018/05/vrouwen-steeds-later-moeder

Chudal, R., Joelsson, P., Gyllenberg, D., Lehti, V., Leivonen, S., Hinkka-Yli-Salomäki, S., ... & Sourander, A. (2015). Parental age and the risk of attention-deficit/hyperactivity disorder: A nationwide, population-based cohort study. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY*, 54(6), 487-494.

Citogroep. (2019). Eindtoets Basisonderwijs. Retrieved from: https://www.cito.com/we-serve/primary-education

Crocetti, E., Branje, S., Rubini, M., Koot, H. M., & Meeus, W. (2017). Identity processes and parent–child and sibling relationships in adolescence: A five wave multi informant longitudinal study. *CHILD DEVELOPMENT,* 88(1), 210-228.

D'Onofrio, B. M., Rickert, M. E., Frans, E., Kuja-Halkola, R., Almqvist, C., Sjölander, A., ... & Lichtenstein, P. (2014). Paternal age at childbearing and offspring psychiatric and academic morbidity. *JAMA PSYCHIATRY,* 71(4), 432-438.

Eastman, M. L., Verhulst, B., Rappaport, L. M., Dirks, M., Sawyers, C., Pine, D. S., ... & Roberson-Nay, R. (2018). Age-Related Differences in the Structure of Genetic and Environmental Contributions to Types of Peer Victimization. *BEHAVIOR GENETICS,* 48(6), 421-431.

Jansen, P. W., Verlinden, M., Dommisse-van Berkel, A., Mieloo, C., van der Ende, J., Veenstra, R., ... & Tiemeier, H. (2012). Prevalence of bullying and victimization among children in early elementary school: Do family and school neighbourhood socioeconomic status matter?. *BMC PUBLIC HEALTH,* 12(1), 494.

Kooijman, M. N., Kruithof, C. J., van Duijn, C. M., Duijts, L., Franco, O. H., van IJzendoorn, M. H., ... & Moll, H. A. (2016). The Generation R Study: design and cohort update 2017. *EUROPEAN JOURNAL OF EPIDEMIOLOGY*, 31(12), 1243-1264.

Lamarche, V., Brendgen, M., Boivin, M., Vitaro, F., Pérusse, D., & Dionne, G. (2006). Do friendships and sibling relationships provide protection against peer victimization in a similar way?. Social Development, 15(3), 373-393.

Lamb, D. J., Middeldorp, C. M., Van Beijsterveldt, C. E., & Boomsma, D. I. (2012). Gene–environment interaction in teacherrated internalizing and externalizing problem behavior in 7 to 12 year old twins. *JOURNAL OF CHILD PSYCHOLOGY AND PSYCHIATRY*, 53(8), 818-825.

Lechner, L., Bolman, C., & Van Dalen, A. (2006). Definite involuntary childlessness: associations between coping, social support and psychological distress. *HUMAN REPRODUCTION,* 22(1), 288-294.

Malaspina, D. (2001). Paternal factors and schizophrenia risk: de novo mutations and imprinting. *SCHIZOPHRENIA BULLETIN,* 27(3), 379-393.

Menesini, E., & Salmivalli, C. (2017). Bullying in schools: the state of knowledge and effective interventions. *PSYCHOLOGY, HEALTH & MEDICINE,* 22(sup1), 240-253.

Merikangas, A. K., Calkins, M. E., Bilker, W. B., Moore, T. M., Gur, R. C., & Gur, R. E. (2017). Parental age and offspring psychopathology in the Philadelphia Neurodevelopmental Cohort. *JOURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY,* 56(5), 391-400.

Merikangas, A. K., Segurado, R., Kelleher, E., Hogan, D., Delaney, C., Gill, M., ... & Heron, E. A. (2016). Parental age, birth order and neurodevelopmental disorders. *MOLECULAR PSYCHIATRY,* 21(6), 728.

Oldehinkel, A. J., Rosmalen, J. G., Buitelaar, J. K., Hoek, H. W., Ormel, J., Raven, D., ... & Hartman, C. A. (2015). Cohort profile update: the tracking adolescents' individual lives survey (TRAILS). *INTERNATIONAL JOURNAL OF EPIDEMIOLOGY,* 44(1), 76-76n.

Oshima, N., Nishida, A., Fukushima, M., Shimodera, S., Kasai, K., Okazaki, Y., & Sasaki, T. (2010). Psychotic like experiences (PLEs) and mental health status in twin and singleton Japanese high school students. *EARLY INTERVENTION IN PSYCHIATRY,* 4(3), 206-213.

Østergaard, S. D., Dalsgaard, S., Faraone, S. V., Munk-Olsen, T., & Laursen, T. M. (2017). Teenage parenthood and birth rates for individuals with and without attention deficit/hyperactivity disorder: a nationwide cohort study. J*OURNAL OF THE AMERICAN ACADEMY OF CHILD & ADOLESCENT PSYCHIATRY,* 56(7), 578-584.

Russell, A. E., Ford, T., Williams, R., & Russell, G. (2016). The association between socioeconomic disadvantage and attention deficit/hyperactivity disorder (ADHD): a systematic review. *CHILD PSYCHIATRY & HUMAN DEVELOPMENT*, 47(3), 440-458.

Salmivalli, C., Kaukiainen, A., & Voeten, M. (2005). Anti bullying intervention: Implementation and outcome. *BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,* 75(3), 465-487.

Schoeler, T., Choi, S. W., Dudbridge, F., Baldwin, J., Duncan, L., Cecil, C. M., ... & Pingault, J. B. (2019). Multi–Polygenic Score Approach to Identifying Individual Vulnerabilities Associated With the Risk of Exposure to Bullying. *JAMA PSYCHIATRY.*

Swagerman, S. C., Van Bergen, E., Dolan, C., de Geus, E. J., Koenis, M. M., Pol, H. E. H., & Boomsma, D. I. (2017). Genetic transmission of reading ability. *BRAIN AND LANGUAGE*, 172, 3-8.

Tearne, J. E., Robinson, M., Jacoby, P., Li, J., Newnham, J., & McLean, N. (2015). Does late childbearing increase the risk for behavioural problems in children? A longitudinal cohort study. P*AEDIATRIC AND PERINATAL EPIDEMIOLOGY,* 29, 41-49. doi: 10.1111/ppe.12165.

Te Velde, E., Habbema, D., Leridon, H., & Eijkemans, M. (2012). The effect of postponement of first motherhood on permanent involuntary childlessness and total fertility rate in six European countries since the 1970s. *HUMAN REPRODUCTION*, 27(4), 1179-1183.

Tully, L. A., Moffitt, T. E., Caspi, A., Taylor, A., Kiernan, H., & Andreou, P. (2004). What effect does classroom separation have on twins' behavior, progress at school, and reading abilities? *TWIN RESEARCH AND HUMAN GENETICS*, 7(2), 115-124.

Van Beijsterveldt, C. E., Groen-Blokhuis, M., Hottenga, J. J., Franić, S., Hudziak, J. J., Lamb, D., ... & Swagerman, S. (2013). The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. *TWIN RESEARCH AND HUMAN GENETICS*, 16(1), 252-267.

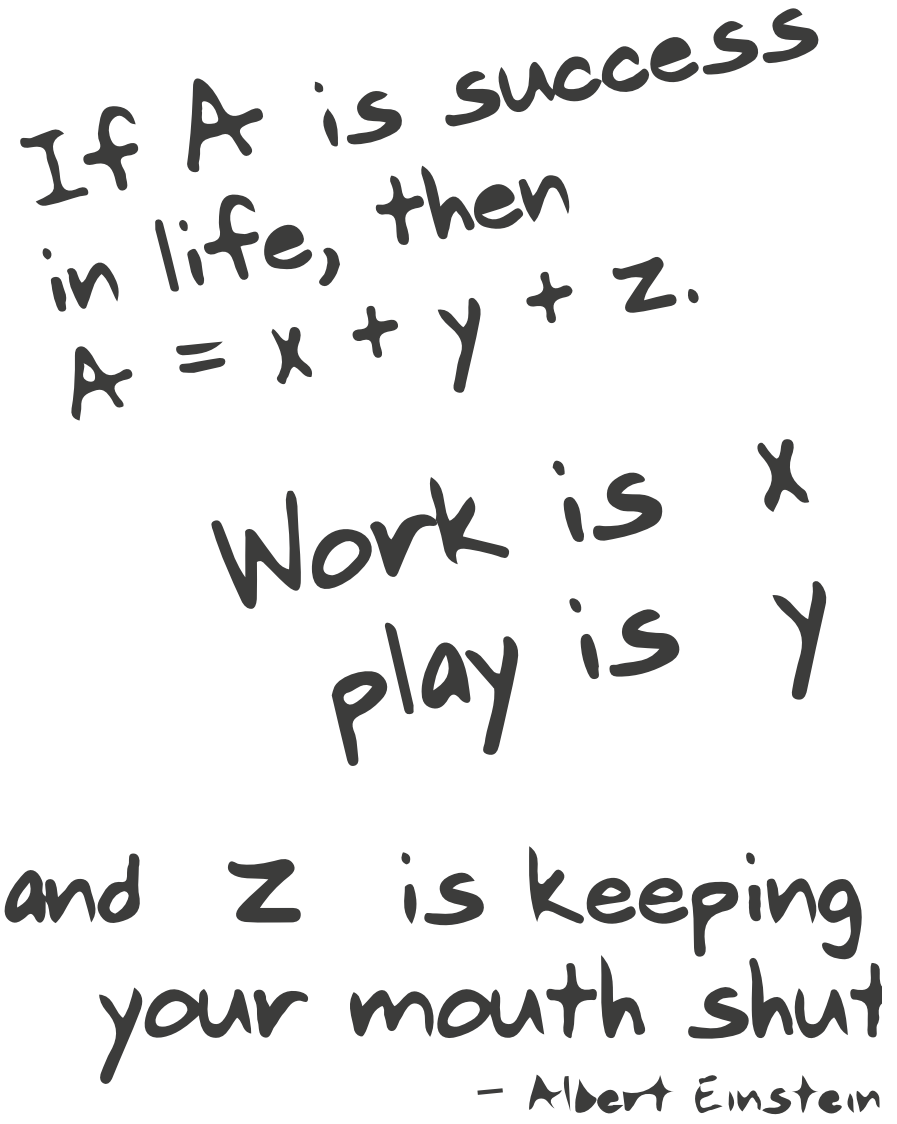
van Bergen, E., van Zuijen, T., Bishop, D., & de Jong, P. F. (2017). Why are home literacy environment and children's reading skills associated? What parental skills reveal. *READING RESEARCH QUARTERLY*, 52(2), 147-160.

Van Leeuwen, M., Van Den Berg, S. M., van Beijsterveldt, T. C., & Boomsma, D. I. (2005). Effects of twin separation in primary school. *TWIN RESEARCH AND HUMAN GENETICS,* 8(4), 384-391.

Vreeman, R. C., & Carroll, A. E. (2007). A systematic review of school-based interventions to prevent bullying. *ARCHIVES OF PEDIATRICS & ADOLESCENT MEDICINE,* 161(1), 78-88.

Weissenberg, R., Landau, R., & Madgar, I. (2007). Older single mothers assisted by sperm donation and their children. *HUMAN REPRODUCTION,* 22(10), 2784-2791.

Dankwoord


If A is success
in life, then
A = x + y + z.
Work is x
play is y
and z is keeping
your mouth shut
- Albert Einstein

Einstein was van mening dat succes een combinatie is van werk, plezier en op sommige tijden je mond dichthouden. Ik zou daar graag een vierde factor aan toevoegen, namelijk ‘steun’. Mijn promotietraject had ik namelijk niet succesvol kunnen afronden zonder alle steun die ik van de mensen in mijn omgeving heb mogen ontvangen. Ik zou nu graag dit stukje papier willen gebruiken om al deze mensen – stuk voor stuk – te bedanken. Ik wil ze bedanken voor alle steun, maar ook voor alle hulp, geduld en gezelligheid.

Steun is echt een bepalende factor geweest in het succesvol afronden van mijn promotietraject, die een hele grote achtbaan bleek te zijn. Voor de mensen die het niet weten: in mei 2015 ben ik begonnen aan dit traject – niet wetende dat ik kort daarna ziek zou worden. Na een leuk, druk en leerzaam eerste jaar merkte ik dat ik steeds vaker ziek werd. In de zomer van 2016 viel ik dan ook definitief uit en kreeg ik te horen dat ik ‘de ziekte van Crohn’ heb. Een hele moeilijke tijd brak aan. Een tijd van vallen, heel diep vallen en weer opstaan. Uiteindelijk kon ik pas eind 2017 weer terugkeren. Helaas heb ik tot het einde toe minder kunnen werken dan ik eigenlijk had gehoopt en heb ik de strijd moeten aangaan met vele obstakels. Des te trotser ben ik dan ook, dat ik ondanks deze moeilijke weg, toch binnen de tijd mijn manuscript heb weten in te leveren. Een mooie quote die daarom erg op mij slaat is:

My illness does not define me. My strength and courage does.

De mensen waar ik mee wil beginnen zijn mijn promotoren, Dorret Boomsma en Meike Bartels, en mijn co-promotoren, Elsje van Bergen en Eveline de Zeeuw. Ik wil jullie allemaal graag bedanken voor jullie goede begeleiding. Meike, met jou heb ik voornamelijk voor mijn eerste project samengewerkt. Ik heb onze afspraken en jouw feedback altijd als zeer prettig ervaren. Ik heb erg veel van jou mogen leren. Dorret, van jou snelle, uitgebreide en ook zeker kritische feedback heb ik veel geleerd. De pittige vragen die jij soms stelde, werkten voor mij zeer goed om in korte tijd veel meer te leren dan anders gelukt zou zijn. Ik kon jou op ieder moment bereiken om advies te vragen om weer verder te kunnen. Elsje, jij was mijn dagelijkse begeleider en ik stapte dan ook vaak eerst op jou af met vragen. Jij was mijn eerste aanspreekpunt. Ik heb ontzettend veel van je geleerd, van Engelse schrijfvaardigheid tot aan het trekken van verbanden. Jij wist met jouw feedback al mijn stukken elke keer nog een stukje beter te krijgen! Onze meetings heb ik altijd als zeer prettig ervaren. Ik ga ze zeker missen! Eveline, met jou heb ik voornamelijk voor de dataverzameling samengewerkt in mijn eerste jaar. Ik heb op dit gebied veel van je geleerd. Ik wil je dan ook ontzettend bedanken voor je begeleiding en al je feedback op mijn projecten.

Extra dank ben ik verschuldigd aan Dorret en Elsje die veel hebben geregeld en mij de kans hebben gegeven om mijn promotie af te kunnen maken ondanks mijn moeilijke situatie. Ik heb geregeld met jullie gesprekken gehad over hoe het ging en samen hebben we gekeken naar de mogelijkheden, hoe ik mijn promotie-traject kon afmaken. Ontzettend bedankt! Ook voor de kansen en vrijheden die jullie mij gegeven hebben om voor het tweede gedeelte van mijn proefschrift zelf het traject op te zetten.

Daarnaast wil ik ook alle tweelingen van het Nederlands Tweelingen Register heel hartelijk bedanken, zonder hun bestaan was mijn proefschrift nooit tot stand gekomen! Ook natuurlijk de ouders en leerkrachten van de tweelingen, bedankt! Zonder hun toestemming en moeite om de vragenlijsten te invullen had ik geen data gehad om te analyseren.

Ook wil ik graag al mijn collega's van BioPsy bedanken. Mijn ex-kamergenootje Jenny, waar ik nog niet eens één dag samen mee op de kamer kon zitten doordat ik ziek werd. Fiona, waar ik gelukkig ook in mijn huidige kamer bij kon zitten. Anne & Yayouk waar ik ook een mooie tijd mee heb gehad in Amerika. Anne, het samen organiseren van het afdelingsuitje vond ik super leuk! Margot, die ik leerde kennen als mijn student-assistent die data ging invoeren voor het NTR en ondertussen zelf ook een AIO is! Natuurlijk ook Hill en Klaasjan. Natascha en Michiel van het secretariaat ook heel hartelijk bedankt! Conor, enorm bedankt voor je advies op statistiek gebied en je bereidheid om te helpen. Toos, bedankt voor het snel in orde maken als ik een vraag had omtrent de NTR data. Ook de overige collega's, bedankt!

Dan nu alle mensen in mijn omgeving die ik graag wil bedanken. Lieve papa en lieve mama, door jullie onvoorwaardelijke steun en (heel erg veel) geduld heb ik het tot hier weten te schoppen. Alle momenten, belangrijk of onbelangrijk, hebben jullie meegekregen en altijd met oprechte interesse geluisterd. Na elk belangrijk gesprek kon ik jullie bellen om alles kwijt te kunnen, wat ik dus ook werkelijk na élk gesprek deed. Jullie zijn ook de reden geweest waardoor ik weer zo snel op het werk kon verschijnen. Terwijl ik nog aardig ziek was, hebben jullie me toch gebracht. Als er ergens een probleem was, of er iets anders moest gebeuren, waren jullie diegene die mij DIRECT hielpen. Lieve papa, ik ben jou eeuwig dankbaar dat je zo'n mooie omslag voor dit proefschrift hebt gemaakt en bovendien de gehele opmaak van het boekje hebt gedaan! Niemand had dit mooier kunnen doen dan jij! Lieve mama, ik ben jou eeuwig dankbaar dat je me meermaals naar de VU hebt gebracht, of uit Utrecht hebt opgehaald als ik daar een afspraak had gehad en toch niet meer zelf thuis kon komen.

Ook hebben jullie mij met alle voorbereidingen geholpen! Jullie weten altijd precies wat ik wil, waardoor ik 100% op jullie kon vertrouwen, zonder zelf alles te moeten checken (wat voor een control-freak als ik soms best lastig is, haha).

En dan mijn lieve vriend, Bas. Ja, je wilt niet in het dankwoord komen te staan, omdat je vindt dat ik alle credits verdien voor dit boekje, maar lieve Bas ook zonder jou lag dit boekje hier nu niet. In mijn laatste jaar van mijn PhD heb ik jou mogen leren kennen, in de meest stressvolle periode van mijn leven. Net chronisch ziek geworden, revaliderend, aan zware medicatie, vaak ziek, druk met PhD en dat terwijl jij zelf ook net uit een hele heftige periode kwam. Alle ingrediënten dus om het samen heel zwaar te krijgen. Dat was het met periodes ook zeker, maar vanaf dat moment werd ook alles beter. Jij hebt er met je zorgzaamheid, je humor en door gewoon te zijn wie je bent, voor gezorgd dat ik steeds meer aan kon. We gingen samen op avontuur, lachten wat af en genoten echt. Dat waren (en zijn nog steeds) super waardevolle momenten, waar ik echt op vooruit kon. Jij beschermde, net als mijn ouders, mij ook door te zorgen dat ik mijn gezondheid in acht bleef nemen. Als ik weer eens eigenwijs was en daar niks op uit deed, kreeg ik een uitbrander. Liters thee en trossen bananen waren jouw specialiteiten, toen je had uitgevonden dat ik daar enorm blij van word. Ook jou moet ik bedanken voor alle ritjes, voor het naar de VU brengen en ophalen uit Utrecht na afspraken. O, wat had je het toen zwaar met mij, want ik schakelde natuurlijk alleen je hulp in als ik écht niet meer kon. Het hielp mij altijd enorm dat ik wist dat als het niet ging, jij ALTIJD zou komen om mij te helpen. Ik kan nog zoveel meer vertellen hoe jij mij gedurende mijn PhD hebt geholpen, maar hier moet ik het maar bij laten. Alleen nog: bedankt dat je mijn paranimf wilt zijn. Het is voor mij een heel fijn idee dat je op deze dag ook echt letterlijk achter me staat!

Ook wil ik de overige onderzoekcohorten TRAILS, RADAR en Gen-R bedanken voor hun fijne samenwerking tijdens onze projecten voor CID. Ik heb zo veel mensen mogen leren kennen en heb van deze samenwerking ontzettend veel geleerd! Bedankt alle co-auteurs! In het bijzonder wil ik Herbert en Mariëlle voor deze samenwerking bedanken, daar heb ik het nauwst mee samen gewerkt. Mariëlle, ik heb het samen begeleiden van het project altijd erg gewaardeerd. Alles konden we overleggen! We vulden elkaar perfect aan. Herbert, bedankt voor de lekkere lunches, waarbij we natuurlijk een goed overleg hadden! Nogmaals Dorret, bedankt voor de vrijheid die ik bij deze projecten gekregen heb, daar heb ik veel begeleidende ervaring mee opgedaan.

Dan wil ik nu mijn zus Daphne en haar man Michiel bedanken. Bedankt voor al jullie interesse en alle foto's van Luna en Emma! Ook die hebben mij er doorheen gesleept. Winnie en Max, ook jullie wil ik graag bedanken. Winnie voor alle interesse, goede zorgen en lekkere kippensoep! Max voor alle interesse en bij tijden het sparren over statistiek. Ook wil ik Tom bedanken dat je achter mij wilt staan, mijn paranimf wilt zijn en ook de sushi-maak-workshoppen had ik niet willen missen! Ook wil ik Bo bedanken. Lief dat je altijd checkte hoe het met mij ging en of je mij kon helpen. Marloes, bedankt voor alle interesse in hoe het met mij gaat en of je nog stukken voor me kon doorlezen! Lisa, bedankt voor alle leuke skype-gesprekken en steun.