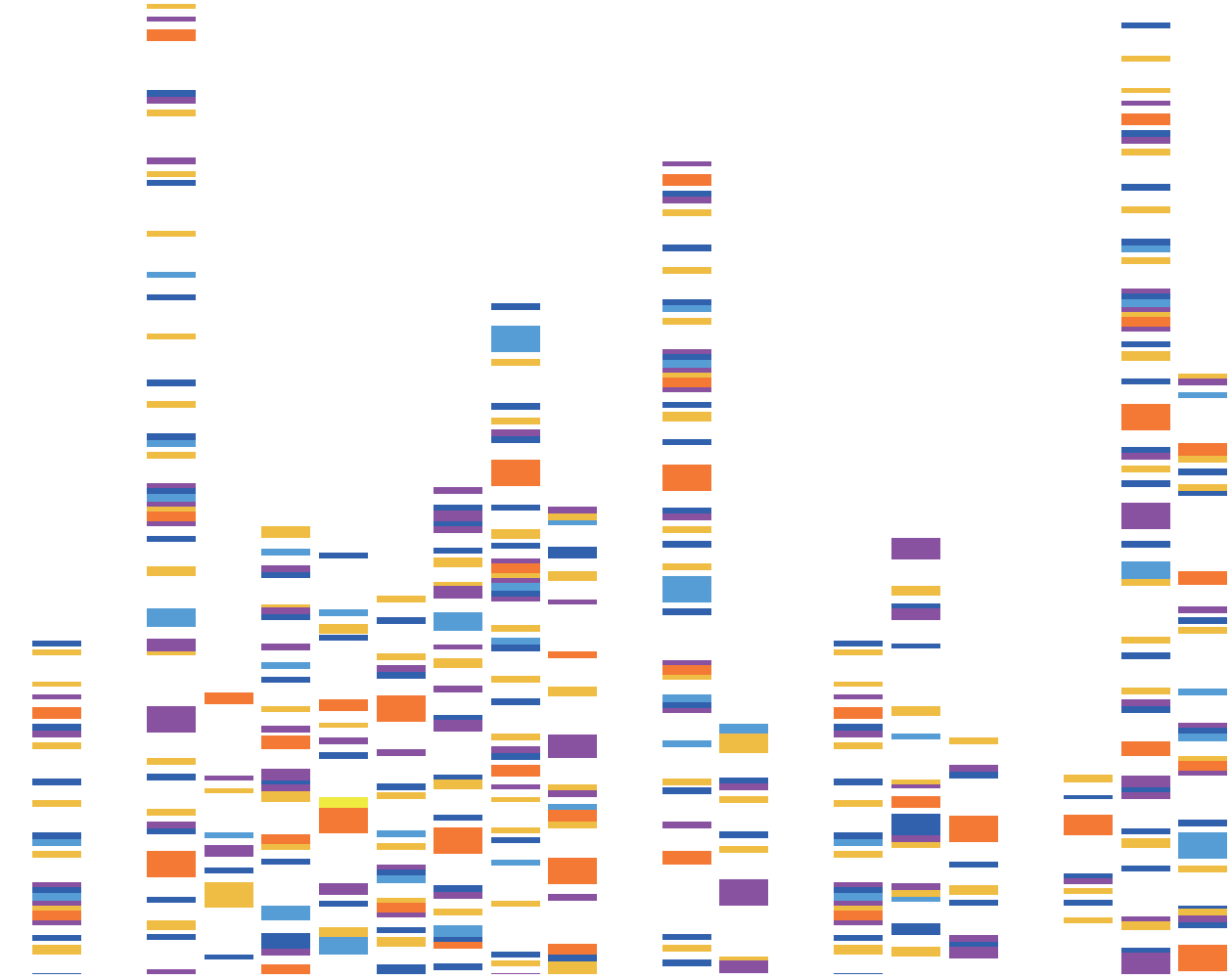
It runs in the family: A genetically informative study of individual differences in aggression

van der Laan, Camiel Morten

2022

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



IT RUNS IN THE FAMILY

A GENETICALLY INFORMATIVE STUDY OF INDIVIDUAL DIFFERENCES IN AGGRESSION

Camiel M. van der Laan

IT RUNS IN THE FAMILY: A GENETICALLY INFORMATIVE STUDY OF INDIVIDUAL DIFFERENCES IN AGGRESSION

Camiel M. van der Laan

Paranymphs: Sjoukje van Deuren
Wonuola Akingbuwa

Layout and print: Guus Gijben, proefschrift-aio.nl

ISBN: 978-94-93270-45-9

VRIJE UNIVERSITEIT

IT RUNS IN THE FAMILY
A GENETICALLY INFORMATIVE STUDY OF INDIVIDUAL DIFFERENCES IN AGGRESSION

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. C.M. van Praag,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op vrijdag 11 maart 2022 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

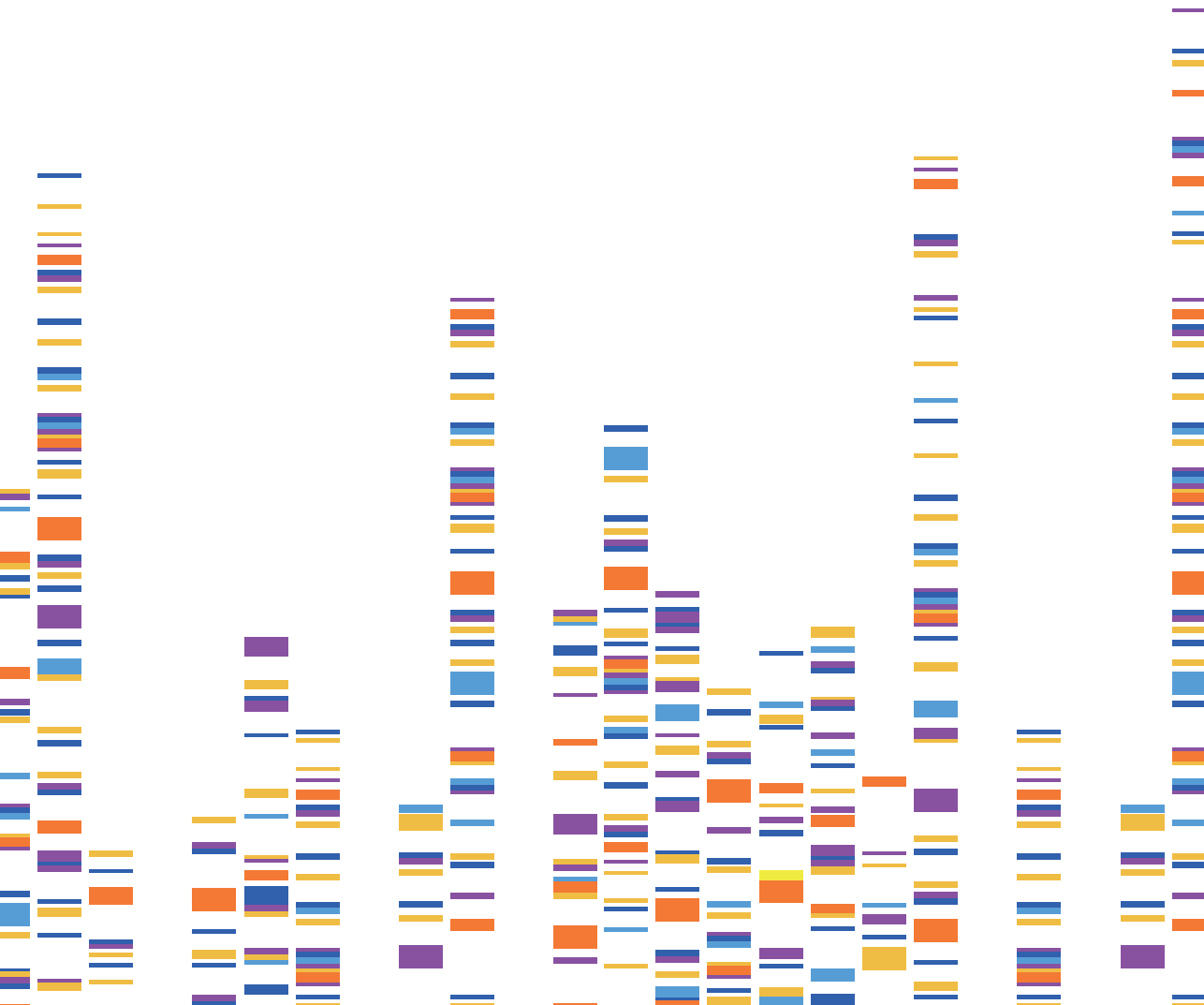
door

Camiel Morten van der Laan

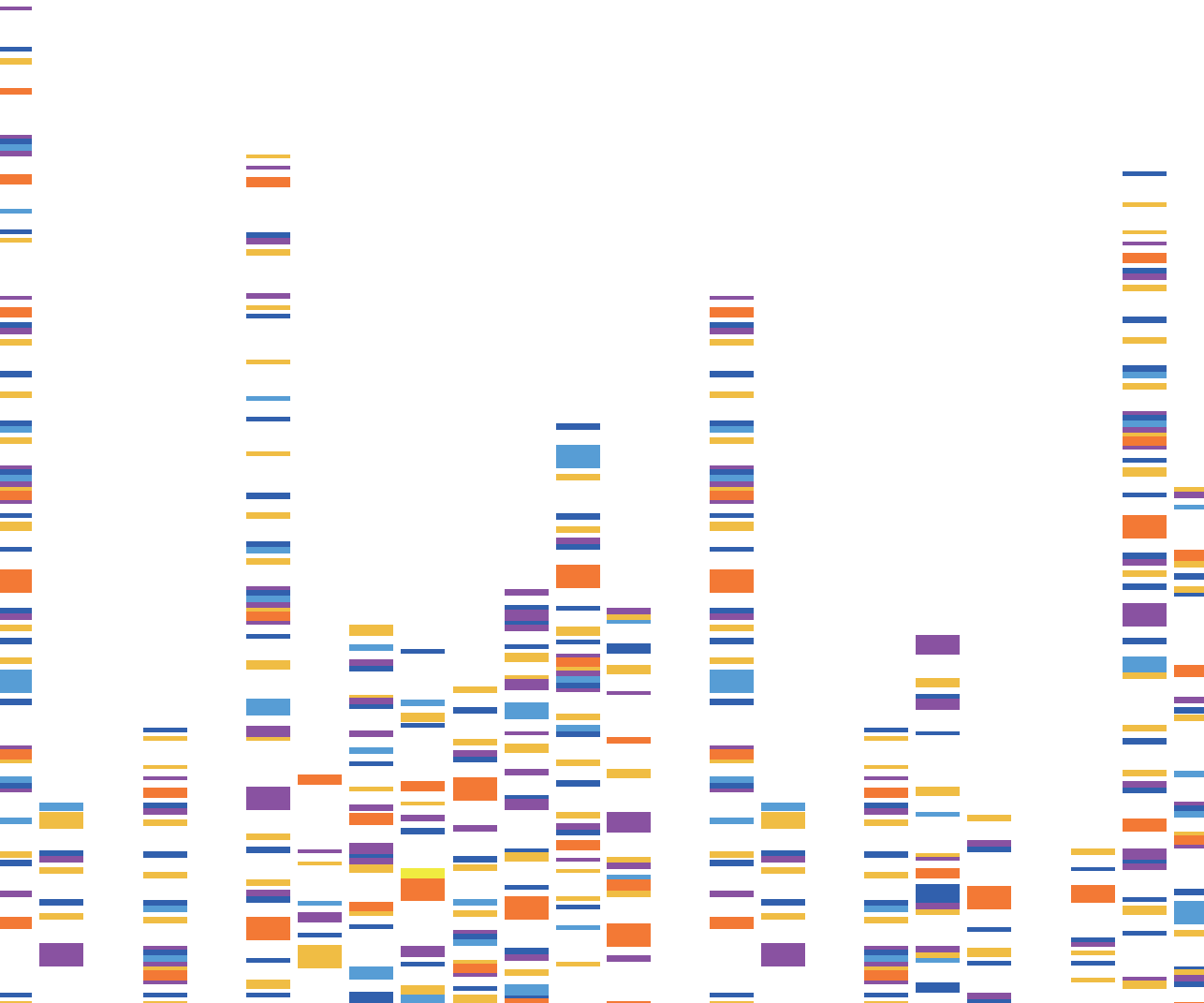
geboren te Leiden

Promotor:	prof.dr. D.I. Boomsma
Copromotoren:	dr. S.G.A. van de Weijer dr. M.G. Nivard
promotiecommissie:	prof.dr. C.V. Dolan dr. C.A. Hartman prof.dr. J.T. Swaab prof.dr. A. Popma dr. V.I. Eichelsheim dr. J.J. Tielbeek

CONTENTS



1
Introduction



INTRODUCTION

Broadly speaking, aggression is behavior that intends to inflict harm on others (Berkowitz, 1993; Lorenz, 1966). Aggression can transpire in various forms, ranging from non-physical, verbal, and relational aggression, to physical aggression (Crick & Grotpeter, 1995, Crick et al., 1999). Typically, physical aggression starts at a young age and peaks at around age two to four years (Alink et al., 2006; Cairns & Cairns 1989, 1994; Karriker-Jaffe et al., 2008; Loeber & Stouthamer-Loeber, 1998; Tremblay et al., 2004; Tremblay, 2010). Physical aggression usually declines as children learn alternative methods to get what they want. Other forms of aggression (e.g. relational, indirect, or social aggression) emerge in the preschool years, continue through childhood and adolescence, and subsequently decline in adulthood (Underwood, 2003). While aggression is relatively common, the propensity for such behavior differs substantively between individuals. Even though aggression typically declines over the life-course, individuals tend to retain their relative position in a population, regardless of their starting position (Pulkkinen & Pitkänen, 1993; Tuvblad & Baker, 2011). In other words, the most aggressive children are more likely to develop into the most aggressive adults (Farrington, 1989). Although the exact causal mechanisms that drive individual differences in aggression and development of aggression over time are still largely unclear, we do know that the propensity for aggression tends to cluster within families (Frisell, Lichtenstein, & Långström, 2011; Margolin et al., 2016, Repetti et al., 2002; van de Weijer et al., 2014; Veroude et al., 2016). In this dissertation the importance of family factors is explored in driving individual differences in aggression, in conforming or desisting from the overall trends we see in aggression over time and across the life-course, and in driving the continuity of aggression across the life-course and across generations.

Interdependence

A wide range of interrelated factors shape our propensity for aggressive behavior, for example genetics, personal experiences, and socio-economic or societal factors. This is reflected in the clustering aggression in families, who share many of these potential causal factors. Investigating these factors is complex, as they are nested in individuals and contexts, which reciprocally influence one another (Bronfenbrenner, 1979). For example, say we are interested in testing for an association between educational attainment and aggression: if we design a study in which we stratify participants on their level of education, we are inherently also stratifying them on their

genotypes because we know that genotype and educational attainment are associated, meaning that any association we find between educational attainment and aggression could be confounded by genotypic influences. To disentangle factors that drive individual differences in aggression, we need an interdisciplinary approach, where we take into account potential confounding effects from correlated factors that may not be of primary interest. A trait such as aggression is of interest in many disciplines, for example in psychology, sociology, medicine, and criminology. Because these disciplines often have different approaches to investigate the influences of a wide range of factors, they each make an important contribution to our understanding of aggression. However, a single disciplinary approach also risks overlooking the complex interrelated structure of potential causal factors. In this light, in Chapter 2, a number of inventive genetically informed research designs are discussed that allow researchers to study the influence of specific factors on aggression while taking into account potential confounding effects.

Aggression is dynamic

When studying factors that potentially shape individual differences in aggression, it is important to realize that aggression is not a static trait. We know that mean aggression levels typically decline across the life-course (Alink et al., 2006; Cairns & Cairns 1989, 1994; Karriker-Jaffe et al., 2008; Loeber & Stouthamer-Loeber, 1998; Tremblay et al., 2004; Tremblay, 2010). We also know that the prevalence of physical fighting and certain extreme forms of aggression (e.g. murder and violent crime) changes over time (FBI, 2019; Frøyland & von Soest, 2018; Kann et al., 2016; Pickett et al., 2013; van der Laan & Goudriaan, 2016), reflecting changes in society. In Chapter 3, the goal is to investigate whether such changes over time are also visible in (generally much less extreme) self-reported aggression. While exploring the effect of time on average aggression levels we ask A) whether changes in aggression over time reflect period or birth-cohort effects, B) whether the overall trends we see in average aggression levels affect everyone equally, and C) whether individual differences in trends in aggression over time cluster within families. The aim here is to better understand the nature of changes in aggression, and to investigate to what extent changes in aggression over time depends on the family you are a part of.

Genetic continuity of aggression

Heritability estimates, i.e. the estimated proportion of variance that can be explained by genetic influences, for aggression and antisocial behavior are on average around 50 percent, and generally slightly increase from childhood through adulthood (Tuvblad and Baker, 2011; Waltes et al., 2016, Odintsova et al., 2019). Different forms of aggression are more or less prevalent at different stages in the life-course. Although individuals retain their genetic make-up throughout their lives, the dynamic nature of aggression across the life-course implies that this does not necessarily mean that the same genetic variants play a role in aggression across the life-course. Longitudinal twin studies on aggression show that genetic factors contribute significantly to the stability of aggression during preschool age, school age, and puberty (van Beijsterveldt et al., 2004; Porsch et al., 2016), and to the stability of externalizing psychopathology in adults (Gustavson et al., 2020).

Ip and colleagues (2021) meta-analyzed genome wide analyses of aggression in 3- to 18 year-olds. This resulted in a total of 328,935 observations from 87,485 unique individuals. Observations were measured across multiple raters, coming from teachers, parents, and self-reports. Despite a lack of single significant hits at the single nucleotide level, Ip et al. (2021) demonstrated that polygenic scores (PGSs), which sum the effects of a range of genetic markers, explained between 0.036 and 0.44 percent of the phenotypic variance in aggression in a hold-out sample of 7 year-old Dutch children (*N*=4,491). In Chapter 4, we test the hypothesis that genetic variants that are associated with aggression during childhood and adolescence are significantly associated with aggression later in the life-course, by modeling the effect of an early life aggression PGS on aggression across the life-course.

Family and aggression

Many risk- and protective factors for the development of aggressive behavior are shared within families and passed on from generation to generation. Based on the differences in genetic similarity between monozygotic and dizygotic twin pairs, we can estimate to what extent differences in behavior can be explained by genetic (the heritability), shared environmental, and unique environmental factors. Shared environmental factors reflect all factors that are not genetic and lead to similarities in outcomes between twins. Unique environmental factors are all factors that are not genetic and lead to differences in outcomes between twins. Implicitly, twin studies can also help us answer questions about what factors drive the clustering

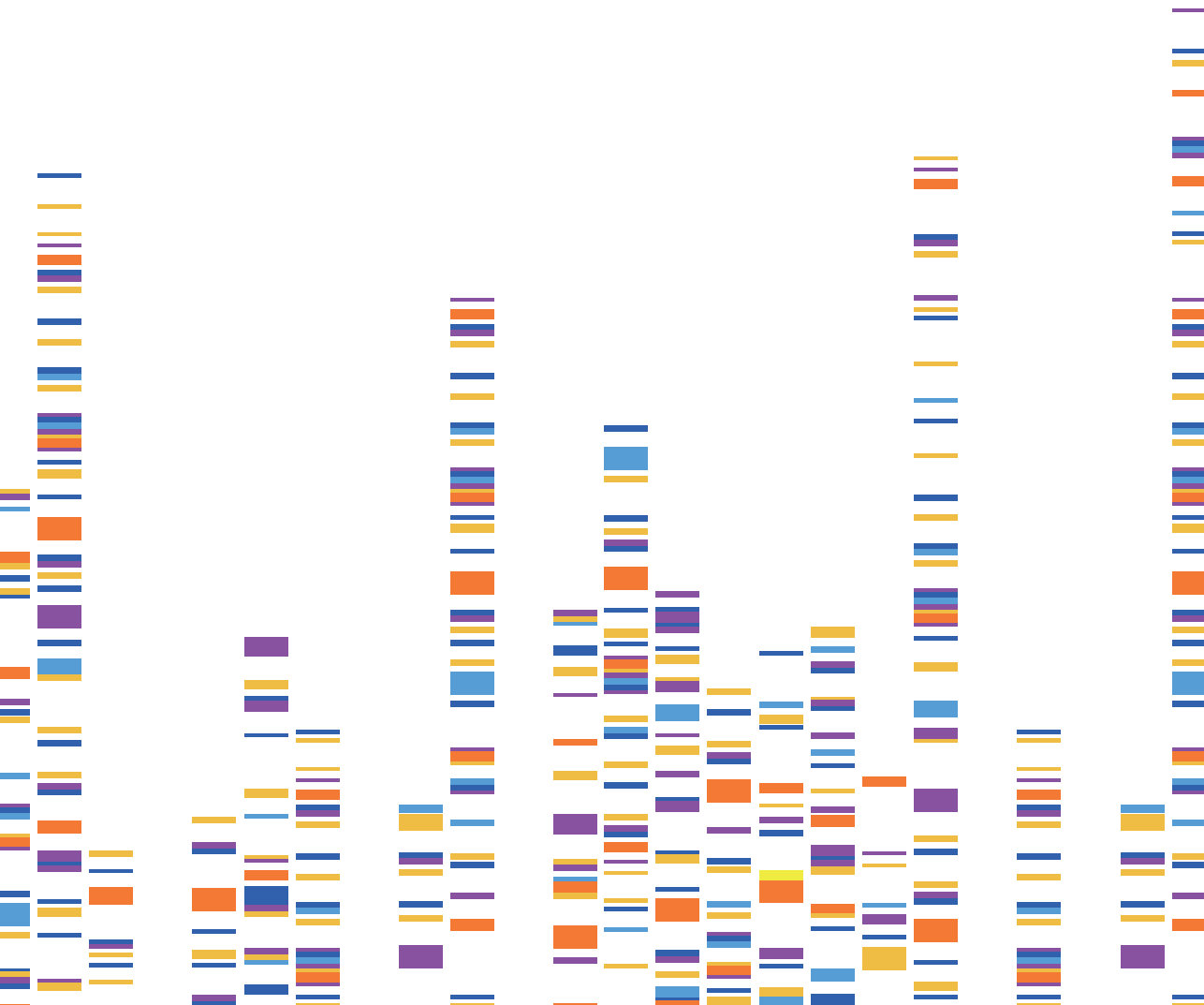
of aggression within families. More specifically, it can tell us whether the intergenerational transmission of aggression is likely driven primarily by genetic transmission, i.e., genetic influences inherently imply genetic transmission of behavior, or cultural transmission, i.e., shared environmental influences are an estimate of the maximum effect of the transmission of cultural norms and values from parents to offspring. In Chapter 5, the goal is to investigate what the relative influences of genetic and cultural transmission are in the familial clustering of rule-breaking behavior, a trait that is closely related to aggression (Bartels et al., 2003).

One important assumption that is often made in behavior genetic studies when estimating heritability, is that genotype and environment are uncorrelated. In practice, this assumption could be violated. For example, the environment in which parents rear their children is partly dependent on their genotypes, i.e., parents with high genetic propensity for aggression are more likely to rear children in an 'aggressive' environment. If offspring behavior is influenced by environment and because offspring inherit their genotype and the environment from their parents, a correlation between the genotype and environment is induced. If a 'passive gene environment correlation' (prGE) is present, genetic effects could reflect correlated environmental effects, and environmental effects may reflect genetic inheritance. One exciting recent development in PGS designs allows to investigate prGE by assessing the effect of non-transmitted parental alleles (i.e., genetic variants) on offspring aggressive behavior (Bates et al., 2018; Kong et al., 2018). In this approach, two PGSs for an individual are calculated, dependent on their parents' genotypes. One PGS for the genetic variants that were inherited from their parents (the transmitted PGS [T-PGS]) and one for the genetic variants that were not inherited (the non-transmitted PGS [NT-PGS]). If the NT-PGS has an effect on offspring aggressive behavior, this effect must be mediated by the environment induced by the genotype of the parents. In chapter 6 the goal is to investigate direct and indirect effects on aggression, and to explore whether genetic transmission (also) reflects indirect effects, or cultural transmission from the genotype of the parents to the phenotype of their children.

Netherlands Twin Register

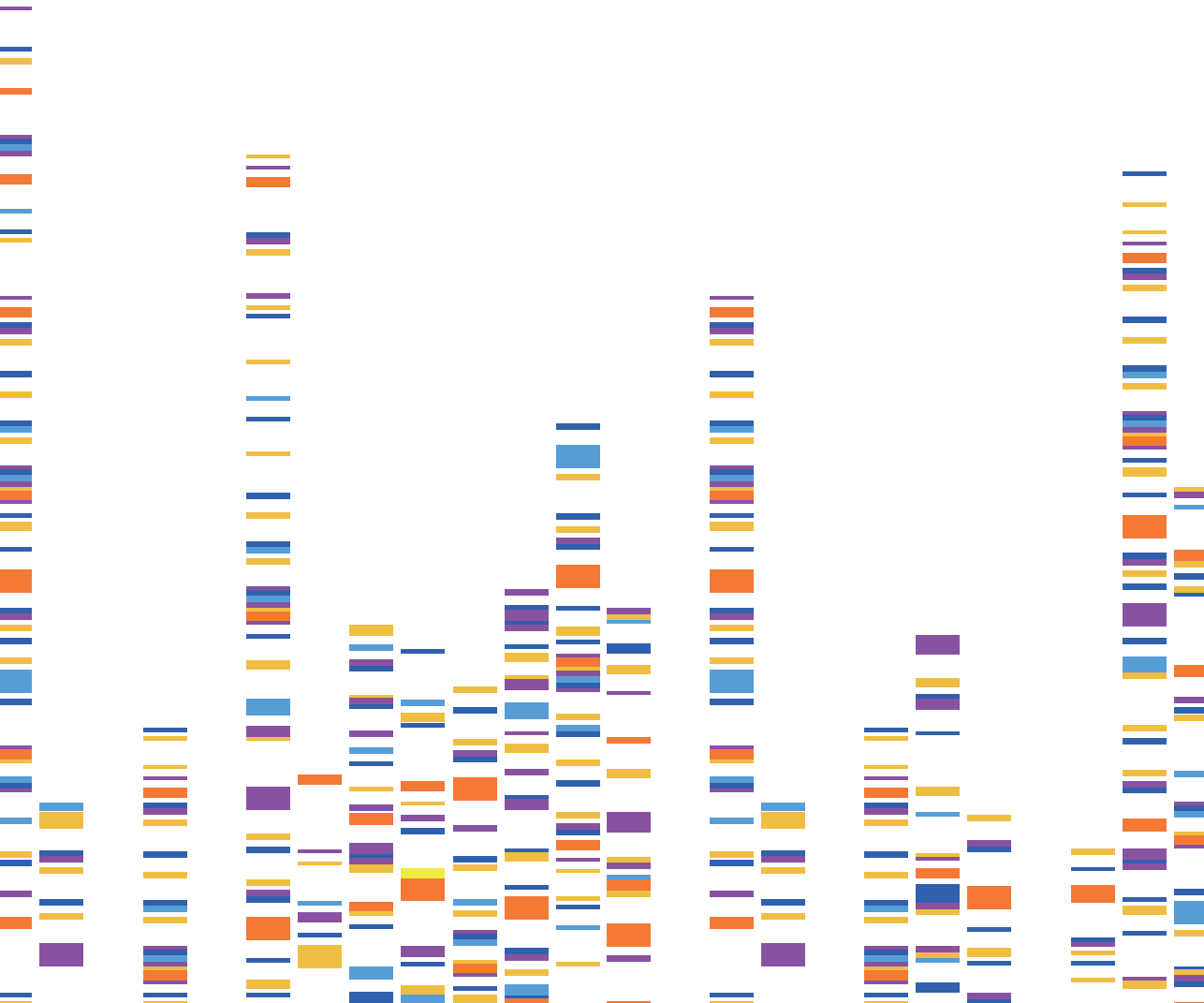
This dissertation would not have been possible without the invaluable contribution of the twins and their family members that are enrolled in the Netherlands Twin Register (NTR). From 1991 onwards, the NTR has collected data on a wide variety of traits (Ligthart et al., 2019). Data on aggression

were collected with multiple measurement instruments by multiple raters: teachers, parents, and self-reports. The majority of the instruments were ASEBA questionnaires: Teacher Report Form, Child Behavior Checklist, Youth Self-Report, and Adult Self-Report (Achenbach & Rescorla, 2000; 2001; 2003). Additionally, a large cohort of twins and siblings were measured with the Devereux Child Behavior Rating Scale (Spivack & Spotts, 1966). In this study I made use of a total of 360,801 aggression observations of 83,027 individuals from 33,653 families. A large proportion of these measures were in twins, but also a large number of their parents, siblings, and spouses completed the NTR surveys. I also analyzed data from families in which parents and their offspring were genotyped across the genome, to define transmitted and non-transmitted PGS. The number of observations for these data are smaller than the number of phenotyped individuals, but the NTR is still unique with respect to the number of families with genotype information on parents and offspring. This thesis would not have been possible without the contribution of these families.



2

Genetische Onderzoekdesigns in de Criminologie: Een 'Toolbox' voor Onderzoek naar Causaliteit en Intergenerationele Transmissie



ABSTRACT

In deze kroniek beargumenteren wij dat sociale en genetische verklaringen voor gedrag niet los van elkaar gezien kunnen worden, en dat onderzoekdesigns die gebruik maken van inzichten uit de genetica een waardevolle toevoeging kunnen zijn aan de 'toolbox' die criminologen voorhanden hebben om onderzoek te doen naar causale verbanden en intergenerationele continuïteit. Vanuit dit perspectief behandelen we vier onderzoekbenaderingen: familie- en tweelingdesigns, genoomwijde associatiestudies, polygenetische scores en Mendelian Randomization. Tot slot bespreken we een aantal praktische overwegingen bij het toepassen van genetische onderzoekdesigns in de criminologie.

Gebaseerd op:
Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2021). Genetische onderzoekdesigns in de criminologie: Een 'toolbox' voor onderzoek naar causaliteit en intergenerationele transmissie. *Tijdschrift Voor Criminologie*, *63*(3), 347-358. https://doi.org/10.5553/TvC/0165182X2021063003004

English summary

In this essay we argue that social and genetic explanations for behavior are intertwined, and that genetically informed research designs can be a valuable addition to the toolbox that is available to criminologists for examining causal relationships and intergenerational transmission. From this 'toolbox' perspective, we discuss four study approaches: twin- and family-designs, genome wide association studies, polygenetic scores and Mendelian Randomization. Finally, we address a number of practical considerations for applying genetically informed research designs in criminology.

Based on:
Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2021). Genetische onderzoekdesigns in de criminologie: Een 'toolbox' voor onderzoek naar causaliteit en intergenerationele transmissie. *Tijdschrift Voor Criminologie*, *63*(3), 347-358. https://doi.org/10.5553/TvC/0165182X2021063003004

GENETISCHE ONDERZOEKDESIGNS IN DE CRIMINOLOGIE: EEN 'TOOLBOX' VOOR ONDERZOEK NAAR CAUSALITEIT EN INTERGENERATIONELE TRANSMISSIE.

Crimineel en antisociaal gedrag zijn complexe menselijke gedragsvormen. Ze worden veroorzaakt door een scala aan sociale en genetische factoren, die vaak niet onafhankelijk van elkaar zijn. Voor een goed begrip van onderliggende mechanismen, de associaties tussen deze factoren en hoe ze zich verhouden tot antisociaal gedrag is een interdisciplinaire aanpak nodig. Binnen de criminologie ligt de nadruk vaak primair op sociale verklaringen voor antisociaal gedrag. Wij beargumenteren in deze kroniek dat sociale en genetische verklaringen voor gedrag niet los van elkaar gezien kunnen worden, en dat genetische onderzoekdesigns een waardevolle toevoeging kunnen zijn aan de toolbox die criminologen voorhanden hebben om causale verbanden en intergenerationele continuïteit te onderzoeken.

In de criminologie ligt vaak de nadruk op sociale verklaringen voor gedrag en worden genetisch sensitieve onderzoekdesigns weinig gebruikt. Genetisch onderzoek laat echter zien dat bijna al het menselijk gedrag een erfelijke component heeft (Polderman e.a., 2015). Erfelijkheid wil zeggen dat verschillen tussen individuen voor een deel te verklaren zijn door genetische verschillen. Menselijk gedrag is polygenetisch en multifactorieel: bijna altijd spelen vele genen een rol, en vrijwel altijd in combinatie met de omgeving. Genetische invloeden zijn dan ook – over het algemeen – niet deterministisch, maar afhankelijk van de omgeving. Dat bijna al het menselijk gedrag erfelijk is, is op zich al interessant, omdat het helpt te begrijpen waarom verschillende mensen in dezelfde omstandigheden verschillend gedrag vertonen. Maar genetisch sensitieve onderzoekdesigns bieden meer mogelijkheden dan het schatten van erfelijkheid. Ook voor criminologen die niet primair geïnteresseerd zijn in genetische verklaringen voor gedrag is er een aantal methodologische en technologische ontwikkelingen in genetische onderzoekdesigns die van waarde kunnen zijn bij onderzoek naar de invloed van sociale factoren op crimineel, agressief en antisociaal gedrag.

De 'gouden standaard' van onderzoek naar causale verbanden, de Randomized Controlled Trial (RCT), is duur en – in de criminologie – vaak niet mogelijk of onethisch. Dit maakt het moeilijk om causaliteit te testen. Observationeel onderzoek kan leiden tot verkeerde interpretaties van associaties, bijvoorbeeld door omgekeerde causaliteit of confounding. Onder de noemer 'toolbox'

bespreken we vier onderzoekdesigns, die gebruik maken van inzichten uit de genetica en een belangrijke toevoeging kunnen zijn bij het testen van causaliteit van robuuste associaties en het vergroten van de kennis van mechanismen achter intergenerationele continuïteit. We beginnen met designs waar de onderzoekspopulatie bestaat uit familieleden, waaronder tweelingen. Vervolgens bespreken we de genoomwijde associatiestudies (GWAS'en) en twee benaderingen die gebruik maken van resultaten uit GWAS'en: polygenetische scores (PGS'en) en Mendelian Randomization (MR).

Tweeling- en familiedesigns

Tweelingonderzoek wordt omschreven als 'the perfect natural experiment to separate familial resemblance from genetic influence' (Martin e.a., 1997, 387). Tweelingstudies maken gebruik van de verschillen in genetische verwantschap tussen eeneiige (monozygote: MZ, met vrijwel identieke genen) en twee-eiige (dizygote: DZ, delen 50 procent van het DNA dat doorgaans varieert tussen mensen) tweelingen om het relatieve belang van genetische factoren (G), gedeelde (C) en unieke omgevingsfactoren (E) te onderscheiden. Gedeelde omgeving, aangeduid met de C van common environment, representeert omgevingsfactoren die leiden tot gelijkenissen in gedrag of andere eigenschappen tussen mensen die opgroeien en leven in dezelfde omgeving, vaak hetzelfde huishouden, zoals gezinsleden. Unieke omgeving, aangeduid met de E van environment, behelst omgevingsinvloeden die leiden tot verschillen tussen familieleden. Voor een complete Nederlandstalige uiteenzetting van het klassieke tweelingmodel verwijzen we naar Van der Laan en collega's (2019).

100 jaar aan tweelingstudies heeft laten zien dat bijna al het menselijk gedrag mede wordt bepaald door erfelijke factoren (Polderman e.a., 2015). Zo wordt bijvoorbeeld gemiddeld 50 procent van de variantie in agressief gedrag verklaard door genetische verschillen tussen mensen (Veroude e.a., 2016; Odintsova e.a., 2019). Soortgelijke schattingen van erfelijkheid werden ook gevonden in meta-analyses naar antisociaal en crimineel gedrag (Ferguson, 2010; Rhee & Waldman, 2002). Dit betekent dat wanneer we alleen naar omgevingsfactoren kijken bij het verklaren van individuele verschillen in antisociaal gedrag, we een incompleet beeld krijgen. Verder valt op dat de geschatte invloed van de gedeelde omgeving in tweelingstudies van antisociaal en crimineel gedrag vaak relatief klein of zelfs afwezig is, wat suggereert dat clustering van probleemgedrag binnen gezinnen voor een groot deel komt door genetische transmissie van ouders op kinderen, en in mindere mate door het doorgeven van normen en waarden binnen

families, oftewel culturele transmissie (Van der Laan e.a., 2019). Bij andere eigenschappen die gerelateerd zijn aan antisociaal en crimineel gedrag, zoals het aantal afgeronde opleidingsjaren (Branigan e.a., 2013), wordt de invloed van de gedeelde omgeving hoger geschat. Gen-omgevinginteractie en -correlatie worden niet altijd expliciet gemodelleerd in tweelingstudies. In de modellen met alleen hoofdeffecten wordt de variantie als gevolg van een interactie tussen A en C onderdeel van genetische variantie en bij een interactie tussen A en E van omgevingsvariantie. Als A en C zijn gecorreleerd, wordt hun covariantie toegeschreven aan de variantie verklaard door C, en als er een correlatie is tussen A en E aan de genetische variantie (Purcell, 2002). Dit betekent dat er bij afwezigheid van C ook geen indicatie is voor een correlatie tussen A en C. Er kan wel nog sprake zijn van een interactie-effect tussen A en C.

De genetische factoren en gedeelde en unieke omgeving in genetische epidemiologie en tweelingstudies kunnen worden gemeten of worden geoperationaliseerd als latente (ongemeten) factoren. Traditioneel was de start van genetisch onderzoek het schatten van de relatieve invloed van genotype en omgeving in tweeling- of familieonderzoek en bestonden vervolgstappen uit het identificeren van de genetische varianten en de specifieke omgevingsfactoren. Een van de problemen die hierbij een rol spelen, is dat variabelen die we wellicht als omgevingsfactoren duiden, ook een belangrijke genetische component kunnen hebben (Vinkhuyzen e.a., 2010). Tegenwoordig worden methoden om genetische varianten te identificeren ook vaak meteen ingezet. Dat geldt met name voor de genoomwijde associatiestudie (GWAS), waarbinnen de associatie met grote aantallen DNA-varianten wordt getest. Recent is een generalisatie van deze methode ook voorgesteld om op eenzelfde wijze, in één model, de associatie met grote aantallen omgevingsfactoren te testen. Maar waar het meten van DNA-varianten door technologische ontwikkelingen steeds makkelijker wordt, geldt dat nog niet voor omgevingsfactoren.

Menselijk gedrag is polygenetisch, vele genen hebben een kleine invloed op gedrag, en vrijwel altijd in combinatie met de omgeving. Eenzelfde gen kan ook van invloed zijn op meerdere uitkomsten, dit heet genetische pleiotropie, en evidentie hiervoor komt uit multivariate genetische modellen. Wanneer er sprake is van genetische effecten op zowel 'omgevingsfactoren' als een uitkomstmaat, zoals antisociaal gedrag, kan er sprake zijn van genetische confounding. Dan wordt wel een associatie geobserveerd, maar is er geen causaal verband.

Het discordante tweelingmodel is een sterke methode om te controleren voor genetische confounding. MZ-tweelingen delen 100 procent van hun erfelijk materiaal. Wanneer tweelingen verschillen in een bepaalde exposure (d.w.z. discordant zijn voor blootstelling aan een risicofactor), kunnen we de associatie tussen de exposure en een uitkomstmaat testen met volledige controle voor genetische confounding en gedeelde omgevingsfactoren, inclusief bijvoorbeeld prenatale factoren. Wanneer een – in de populatie – robuuste associatie wegvalt of minder sterk is in discordante MZ-tweelingparen, duidt dit op confounding. Dit kan genetische confounding zijn: er is dan een pleiotropisch effect waarbij genetische factoren de uitkomst en de exposure beïnvloeden. Of er is confounding doordat gedeelde omgevingsfactoren zowel de exposure als de uitkomst beïnvloeden. Caspi en collega's (2004) lieten bijvoorbeeld zien dat de associatie tussen emotionele expressie van de moeder en antisociaal gedrag in kinderen overeind bleef in een discordant tweelingmodel, en interpreteren hun resultaten als een ondersteuning voor een causaal verband.

Niet voor iedere vraagstelling zijn data van discordante MZ-tweelingparen beschikbaar. Een goed alternatief is dan het discordante siblingmodel, waarbij niet-blootgestelde broers en zussen uit hetzelfde gezin als controlegroep fungeren. Zo wordt deels voor de genetische verschillen (broers en zussen delen gemiddeld 50 procent van hun genen) gecontroleerd, en voor gedeelde omgevingsfactoren. Verschillende studies hebben zo laten zien dat verbanden minder sterk worden in een 'within'-familiedesign. Sariaslan en collega's (2013) toonden bijvoorbeeld aan dat het leven in een achterstandswijk onder Zweedse adolescenten geassocieerd was met een 57 procent hoger risico op het plegen van een geweldsdelict. Dit significante verband verdween echter volledig in een discordant siblingmodel, waar de broers en zussen in een andere wijk woonden en naar andere scholen gingen. Een Nederlandse studie liet zien dat (vroeg) crimineel gedrag samenhangt met een 11 procent hogere kans op vroegtijdig schoolverlaten, maar dat dit daalt naar 5 procent in discordante broers en zussen, en naar slechts 3 procent in discordante tweelingen van hetzelfde geslacht (Rud e.a., 2018). Voor zowel het discordante tweelingmodel als het discordante siblingdesign geldt dat het overeind blijven van een associatie een voorwaarde is voor causaliteit, maar het is geen bewijs voor causaliteit. Er kan nog steeds confounding zijn door niet-gedeelde omgevingsfactoren; met name wanneer beide eigenschappen op hetzelfde moment gemeten zijn, kan de richting van het causale verband niet duidelijk zijn.

Naast het discordante tweeling- of siblingdesign is het soms ook mogelijk om hypotheses over causale verbanden te testen in bivariate tweeling- en familiemodellen, waarbij alle data worden geanalyseerd (niet alleen de data van discordante paren). Een vereiste hiervoor is dat de twee variabelen van elkaar verschillen in hun 'genetische architectuur'. Als bijvoorbeeld variabele A hoog erfelijk is en variabele B voornamelijk wordt beïnvloed door gedeelde omgevingsfactoren, dan kan worden getoetst of een model waar variabele A een invloed heeft op variabele B beter past dan het alternatief waar variabele B een invloed heeft op variabele A. Deze 'Direction of Causation' (DoC)-modellen (Heath e.a., 1993; Duffy & Martin, 1994) zijn bijvoorbeeld toegepast in een recente studie (Willems e.a., 2020), waarin bleek dat familieconflict waarschijnlijk leidt tot lagere zelfcontrole, en niet andersom.

Genoomwijde associatiestudies

Een benadering die de afgelopen twee decennia veel aan momentum heeft gewonnen, is de genoomwijde associatiestudie (GWAS). In een GWAS wordt onderzocht of een genetische variant, meestal een single-nucleotide polymorfisme (SNP), is geassocieerd met een bepaalde – continu of dichotoom gemeten – eigenschap. Genoomwijd wil zeggen dat dit voor honderdduizenden of miljoenen SNP's wordt gedaan. Deze SNP's liggen op bekende locaties in het genoom waar personen van elkaar kunnen verschillen.

GWAS'en zijn tegenwoordig groot. Neem bijvoorbeeld een recent onderzoek naar het aantal afgeronde opleidingsjaren waar ongeveer 7,1 miljoen SNP's zijn onderzocht onder ruim 1,1 miljoen proefpersonen (Lee e.a., 2018). Omdat veel eigenschappen een sterk polygenetische achtergrond hebben – veel genetische varianten spelen een rol, die ieder een klein effect hebben op een eigenschap –, zijn grote steekproeven nodig om deze kleine effecten te kunnen detecteren. Daarnaast is de statistische power ook laag vanwege de enorme aantallen associaties die getest worden. Over het algemeen wordt een genoomwijd significantieniveau van $p<5x10-8$ ($p<0,00000005$) aangehouden, wat neerkomt op een Bonferroni-correctie voor 1 miljoen onafhankelijke tests. Dit aantal is gebaseerd op de schatting van het aantal onafhankelijke stukken informatie in het menselijk genoom (Pe'er e.a., 2008).

Recente GWA-onderzoeken naar antisociaal gedrag (N=31.968; Tielbeek e.a., 2018) en agressief gedrag in kinderen (N=87.485; Ip e.a., 2021) maakten nog gebruik van relatief kleine steekproeven, en bij de primaire analyses werden bij beide geen significante associaties gevonden. Ter vergelijking: de

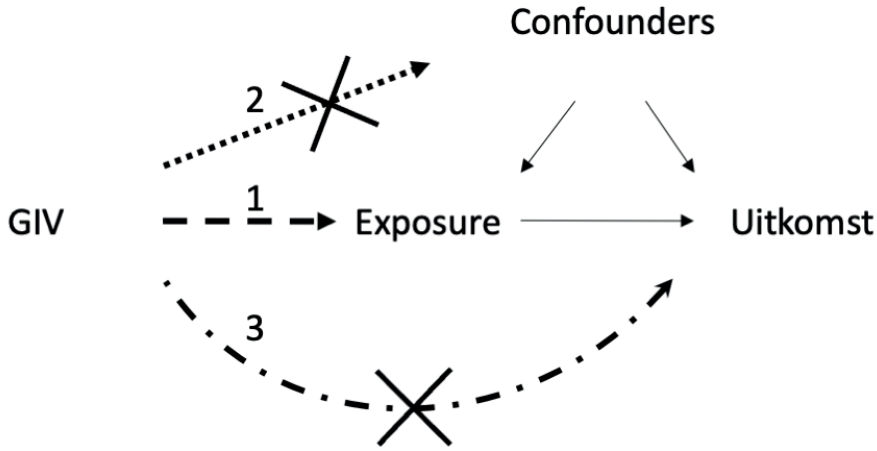
eerdergenoemde GWAS naar het aantal afgeronde opleidingsjaren, met een steekproef van ruim 1,1 miljoen personen, vond 1.271 unieke significante associaties (Lee e.a., 2018). De geschatte erfelijkheid van afgeronde opleidingsjaren uit tweelingstudies is ongeveer 40 procent (Branigan e.a., 2013). Dat is iets lager dan bijvoorbeeld bij agressief en antisociaal gedrag (ongeveer 50 procent; Ferguson, 2010; Odintsova e.a., 2019). We verwachten dan ook dat GWA-onderzoek naar agressief en antisociaal gedrag vergelijkbare of zelfs hogere aantallen significante associaties oplevert in grotere steekproeven. Er worden nu methoden ontwikkeld om de steekproefomvang te vergroten door bijvoorbeeld verschillende indicatoren van agressief of externaliserend gedrag tegelijk te analyseren in een meta-analyse (Ip e.a., 2021; Linner e.a., 2021).

Bijna twintig jaar GWAS heeft een enorme hoeveelheid resultaten opgeleverd (Buniello e.a., 2019; Visscher e.a., 2012, 2017). Deze resultaten zijn nuttig voor vervolgonderzoek. Zo kan met de resultaten worden gekeken naar de genetische correlaties tussen eigenschappen, zelfs als deze niet in dezelfde steekproef zijn gemeten. Antisociaal gedrag is bijvoorbeeld genetisch gecorreleerd met cannabisgebruik (rg=0,69) en aantal sigaretten per dag (rg=0,59), maar niet met wekelijks alcoholgebruik en of iemand ooit gerookt heeft (Tielbeek e.a., 2018). Genetische correlaties tussen eigenschappen betekenen dat associaties een gedeelde genetische achtergrond hebben. Dat kan ontstaan door pleiotropie, of wanneer de ene variabele de andere beïnvloedt. Daarom kunnen de resultaten uit een GWAS ook een 'instrument' voor onderzoek naar causaliteit worden in Mendelian Randomization (MR)-onderzoek.

Mendelian Randomization

Mendelian Randomization (zie Figuur 2.1) is een veelbelovend alternatief voor de RCT, dat mogelijk ook in de criminologie kan worden gebruikt. Bij MR wordt vaak informatie over één of meer SNP's uit GWAS als genetische instrumentele variabele (GIV) gebruikt. Voor toepassing van MR moet aan een drietal vrij strikte voorwaarden voldaan worden (Smith & Ebrahim, 2003; Boutwell & Adams, 2020). Ten eerste moet de GIV (robuust) geassocieerd zijn met de exposure (de mogelijke causale variabele) waarin we geïnteresseerd zijn. Ten tweede moet de GIV onafhankelijk zijn van confounders van de exposure-uitkomstassociatie. En tot slot mag de GIV niet direct geassocieerd zijn met de uitkomst.

Individuen erven willekeurig, en onafhankelijk van de omgeving, hun genetische varianten van hun ouders. Die genetische varianten liggen vast vanaf conceptie, en gaan daarom temporaal altijd vooraf aan de eigenschappen en gedragingen die we onderzoeken. Als aan de drie voorwaarden wordt voldaan, is bij een causaal verband de invloed van de GIV op de uitkomst indirect, en geheel gemedieerd door de invloed op de exposure.



Figuur 2.1. Het Mendelian Randomization (MR) model

Noot: Een causaal verband wordt getest door de associatie tussen een genetische instrumentele variabele (GIV) (die invloed heeft op de mogelijk causale variabele) en een uitkomst te schatten. De genummerde lijnen geven de drie assumpties weer. Lijn 1: een robuuste associatie tussen de GIV en de exposure. Lijn 2: de GIV moet onafhankelijk zijn van confounders van de exposure-uitkomstassociatie. Lijn 3: de GIV mag niet direct geassocieerd zijn met de uitkomst.

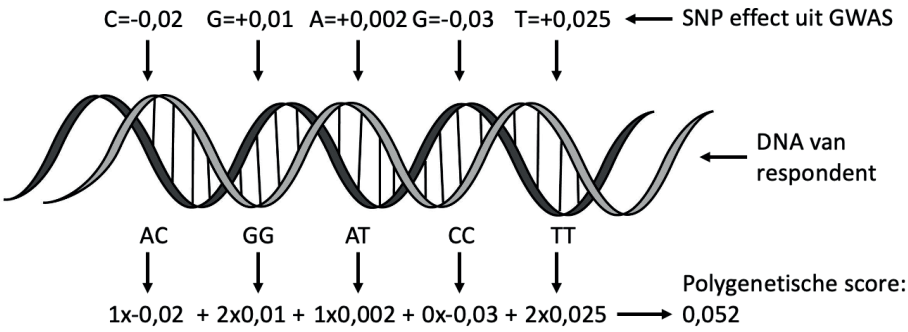
Op dit moment kunnen we alleen de eerste voorwaarde direct testen. Door de toename van grootschalige GWAS stijgt het aantal genetische varianten die significant geassocieerd zijn met verschillende exposures, en zo als GIV kunnen dienen. Associaties tussen de GIV en confounders van de exposure-uitkomstassociatie kunnen zich voordoen door toeval, door correlaties tussen de GIV en andere genetische varianten, en door pleiotropie: de GIV beïnvloedt zowel de exposure als de confounders. Recente ontwikkelingen in MR-onderzoek zijn onder meer de combinatie met een 'within-family'-design. Brumpton en collega's (2020) laten zien dat een causaal effect van BMI en lengte op het aantal afgeronde opleidingsjaren grotendeels verdwijnt in zo'n design. GIV's zijn niet beperkt tot een enkele genetische variant. Dudbridge (2021) laat zien hoe ook PGS'en als GIV kunnen worden gebruikt. Omdat hiermee veel meer SNP's tegelijk als GIV worden gebruikt, wordt de verklaarde variantie in de exposure groter, wat leidt tot meer statistische

power. Adams en Boutwell (2020) onderzochten met MR met PGS de relatie tussen avontuurlijkheid en tien cognitieve en gedragskenmerken. Ze concludeerden dat avontuurlijkheid aan twee kanten snijdt: zo leidt dit bijvoorbeeld tot een hoger opleidingsniveau, maar ook tot een hogere mate van risicogedrag, zoals te hard rijden.

Polygenetische score (PGS)

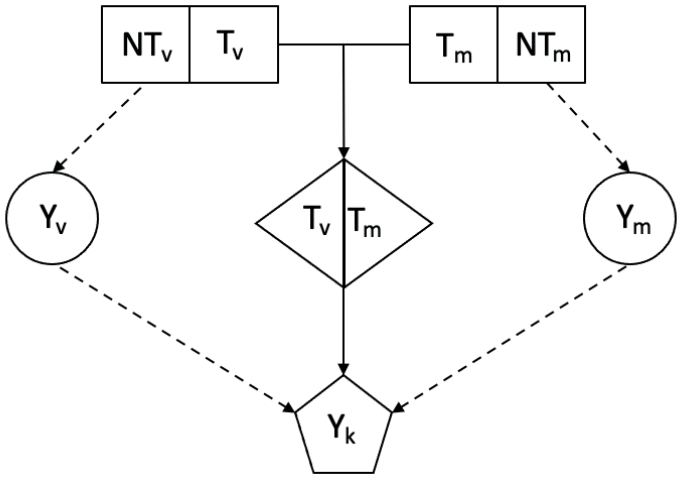
Met de resultaten uit GWA-onderzoek kunnen polygenetische scores (PGS'en) worden geconstrueerd (Choi e.a., 2020), waarbij GWA-resultaten worden terugvertaald naar individueel niveau. In een GWAS wordt het effect van een SNP op een uitkomst geschat. Met deze kennis kunnen we voor elke persoon schatten wat zijn genetische predispositie is voor een bepaalde eigenschap. Bijvoorbeeld: een persoon heeft genotype AA, een tweede persoon AT en een derde TT. Als A de uitkomst van een continue variabele met 0,002 verhoogt, dan scoort de eerste persoon 0,004, de tweede 0,002 en de derde 0 (zie Figuur 2.2). Bij een PGS wordt gesommeerd over het hele genoom (alle onderzochte SNP's), of bijvoorbeeld alleen over de effecten van significant geassocieerde SNP's. PGS'en kunnen voor tal van onderzoeksvragen worden gebruikt, zoals predictie op basis van de genetische predispositie van een persoon. Een PGS van agressief gedrag in kinderen verklaart op het moment echter slechts 0,4 procent van de variantie in agressief gedrag in een steekproef van 7-jarigen (Ip e.a., 2021). Maar zelfs als predictie op individueel niveau niet mogelijk is, kan PGS-analyse helpen bij het begrijpen van bepaalde fenomenen in de criminologie. Zo onderzochten wij in een recente studie of PGS'en die geassocieerd waren met agressie in kinderen op groepsniveau geassocieerd waren met agressie in adolescenten en volwassenen. De resultaten lieten zien dat dezelfde genetische varianten ook op latere leeftijd nog invloed hebben, wat voor een deel verklaart waarom individuele verschillen in agressief gedrag gedurende de levensloop vaak stabiel zijn (Van der Laan e.a., 2021-c).

PGS'en spelen ook een belangrijke rol in onderzoek naar intergenerationele transmissie. Kinderen erven 50 procent van hun genoom van moeder en 50 procent van vader. Daarnaast groeien zij op in een omgeving die (deels) door de ouders is gecreëerd of geselecteerd. Dit proces van creatie en selectie van omgeving wordt deels weer beïnvloed door het genoom van de ouders. De kinderen erven dus zowel een genetische predispositie voor bijvoorbeeld cognitieve vaardigheden als een omgeving, die op grond van het ouderlijke genoom die vaardigheden stimuleert. De uitdaging voor onderzoekers is om de directe genetische transmissie te onderscheiden van de 'culturele'



Figuur 2.2. Voorbeeld van PGS constructie in een persoon

Noot: A, C, G en T zijn de nucleobasen in het DNA. Op de locaties in het DNA waar personen van elkaar kunnen verschillen (SNP's), heeft een persoon twee nucleobasen. De effecten (geschat in een GWAS) van de aanwezige nucleobasen in het DNA van een persoon worden in een PGS gesommeerd tot een polygenetische score.



Figuur 2.3. Transmitted/ non-transmitted alleles design

Noot: De doorgetrokken lijnen geven het effect aan van de doorgegeven PGS (T) van beide ouders (v=vader, m=moeder) op een eigenschap van het kind (Yk). De onderbroken lijnen geven het effect aan van de niet-doorgegeven PGS (NT), naar de eigenschappen van de ouders (Yv en Ym), naar de eigenschap van het kind.

transmissie. Een van de spannendste ontwikkelingen in PGS-onderzoek is het 'transmitted/non-transmitted alleles'-design (Figuur 2.3; Bates e.a., 2018; Kong e.a., 2018). Hiermee wordt geschat wat het effect is van culturele transmissie van ouder op kind, zonder dat er sprake is van confounding door genetische transmissie. Dit wordt gedaan door te onderzoeken of de

genetische achtergrond van ouders via de omgeving invloed heeft op het gedrag van hun kinderen. Ouders geven de helft van hun DNA door aan hun kinderen en de andere helft niet. Voor beide delen kan een PGS worden berekend. Wanneer er een effect is van de PGS die niet is doorgegeven aan het kind op het gedrag van het kind, dan moet dit effect via de omgeving lopen. Dat het effect van de niet-doorgegeven PGS van ouders via de omgeving een belangrijke rol kan spelen bij de ontwikkeling van kinderen, werd bijvoorbeeld al aangetoond bij het aantal afgeronde opleidingsjaren. Het effect van de niet-doorgegeven PGS werd geschat op 29,9 procent van het effect van de doorgegeven PGS (Kong e.a., 2018). Voor toepassing van het 'transmitted/non-transmitted alleles'-design zijn DNA-data nodig van ouders en kind, en informatie over de eigenschap van het kind die onderzocht wordt.

Epiloog

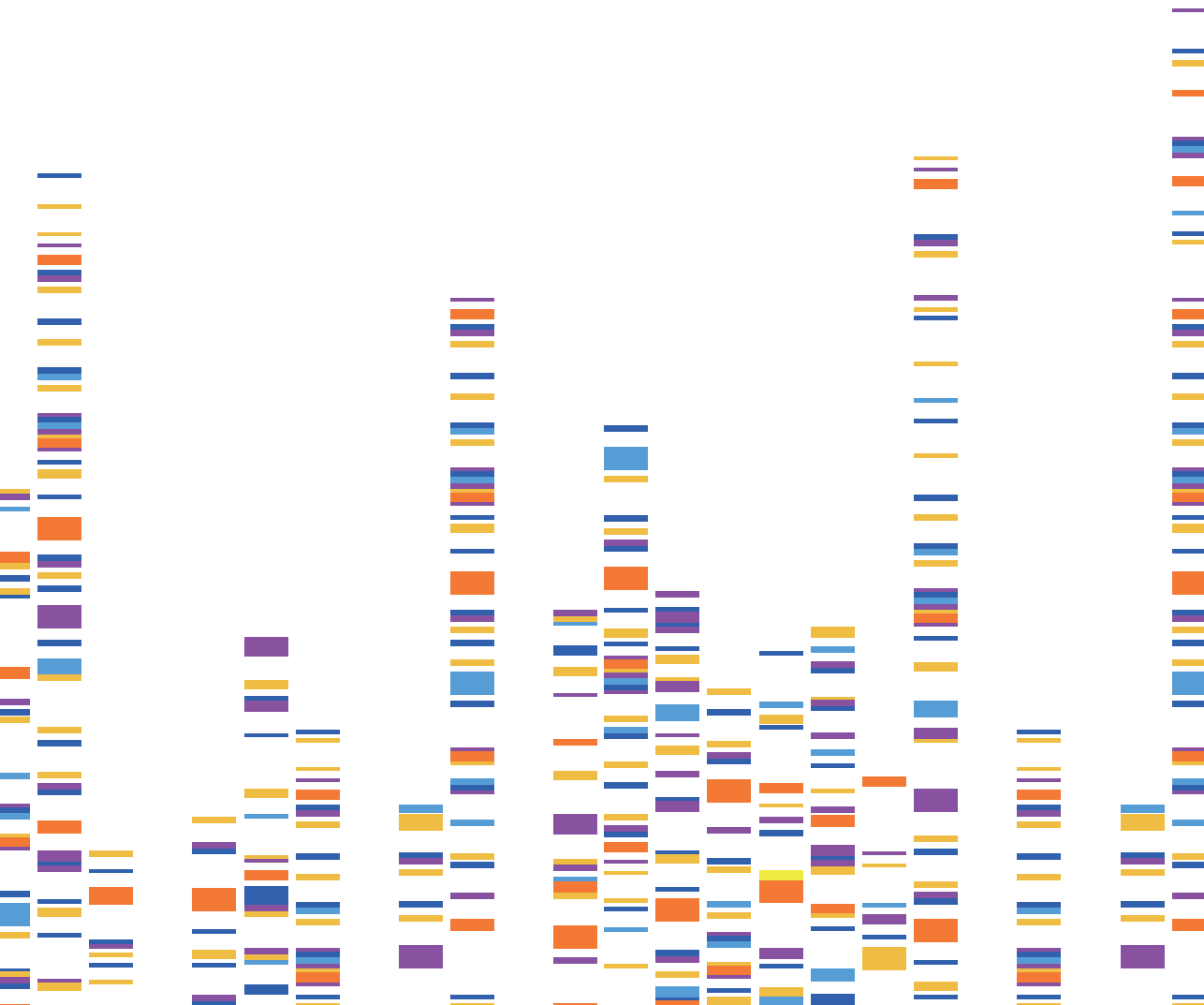
In deze kroniek hebben wij vier onderzoekdesigns besproken die veelbelovend zijn in de epidemiologie en sociale wetenschappen. Deze designs kunnen mogelijk ook gebruikt worden bij het ontrafelen van causale verbanden achter associaties in de criminologie, en bij het beter begrijpen van de mechanismen achter intergenerationele continuïteit van antisociaal gedrag. Wij realiseren ons dat enkel de kennis van de designs niet altijd voldoende kan zijn om deze onderzoeken uit te voeren. Daarom benoemen wij hierna kort de praktische kant van de genetische onderzoekdesigns met een focus op de beschikbaarheid van data.

Met de huidige mogelijkheden is een van de meest praktische manieren om te controleren voor genetische confounding en ongemeten gedeelde omgeving het doen van onderzoek op het niveau van gezinnen en families. Vrijwel elk kwantitatief onderzoekdesign is geschikt om 'within'- en 'between'-familieanalyses uit te voeren. Onderzoeken die gebruik maken van registerdata, zoals de CBS-microdata, zijn hier bij uitstek geschikt voor, omdat hieruit informatie over familieverbanden kan worden gehaald, zoals in het onderzoek van Rud en collega's (2018).

Bij het doen van een tweelingstudie volstaan alleen CBS-data niet altijd. Hiervoor is bijvoorbeeld informatie over zygositeit (een- of twee-eiige tweeling) nodig. Voor die informatie zijn we aangewezen op andere partijen. In Nederland zijn meerdere onderzoeksgroepen die data verzamelen van twee- en meerlingen waarbij zygositeit wordt bepaald, de grootste hiervan is het Nederlands Tweelingen Register (NTR; Ligthart e.a., 2019).

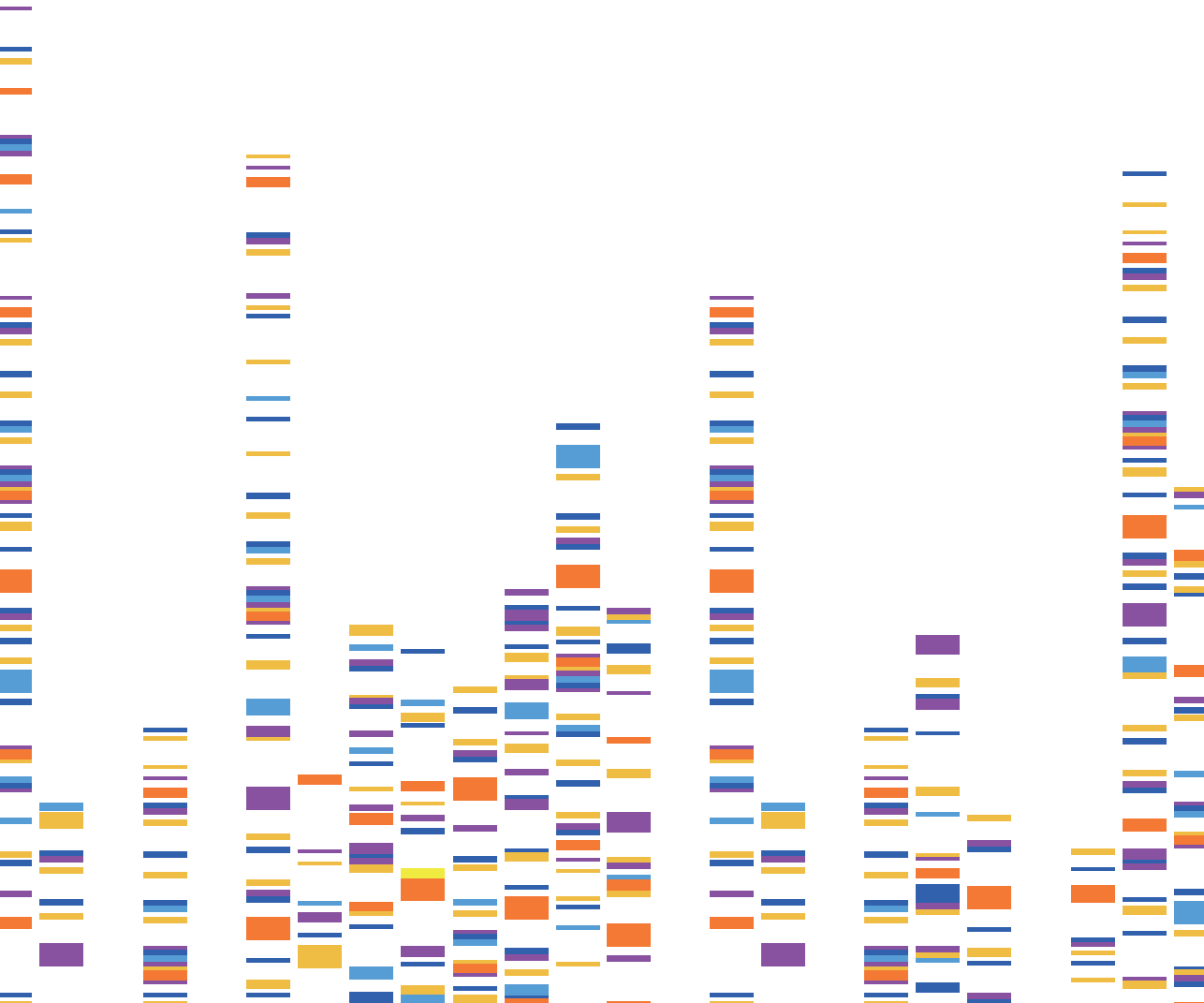
Voor een GWAS, MR-studie of PGS-studie zijn genotypedata nodig. In Nederland zijn enkele honderden biobanken en onderzoeksgroepen die genotypedata verzamelen. Procedures voor het aanvragen van data bij cohorten en biobanken zijn volop in ontwikkeling, net zoals verkenningen naar de mogelijkheden voor 'record linkage' met registergegevens (voorbeelden zijn BBMRI: www.bbmri.nl, en ODISSEI: www.odissei-data.nl). Door de mogelijk sensitieve aard van zowel de cohortdata als de registerdata heeft ODISSEI samen met de Samenwerkende Universitaire Rekenfaciliteiten (SURF) het ODISSEI Secure Supercomputer-platform opgezet. De Zeeuw en collega's (2021) lieten al zien dat NTR-data succesvol en veilig aan CBS-data gelinkt konden worden om een GWAS uit te voeren. Deze ontwikkeling betekent dat het op korte termijn wellicht ook mogelijk is om genetische onderzoekdesigns toe te passen met geregistreerde criminaliteit als uitkomstmaat.

Ter conclusie: de oorzaken van individuele verschillen in gedrag zijn uitermate complex. We weten dat bij bijna alle menselijke eigenschappen en gedragingen zowel genetische invloeden als omgevingsinvloeden een rol spelen. Om deze te scheiden, zijn ingenieuze benaderingen nodig die een beter begrip geven van causale processen. Genetische onderzoekdesigns kunnen hier een waardevolle rol in spelen.



3

Familial Clustering of Trends in Aggression



ABSTRACT

Objectives: Examine trends in aggressive behavior from 1991 to 2015, investigate whether these trends apply equally to all individuals, and explore the extent to which differences in trends over time cluster within families.
Methods: Our study included 69,465 measures from 40,400 individuals, from 15,437 Dutch families. Aggression was measured between 1 and 4 times by self-report. We fitted a mixed effects model, modeling the effect of time, age, and gender on aggression, and considering the three levels of nesting in the data, i.e. repeated measures, individuals, and families. To investigate if individual differences in trends in aggression over time cluster within families, variance in aggression and in time and age effects was partitioned into within- and between family variance components.
Results: We found a steady decline in aggression over time, between 1991 and 2015, as well as over the life course. Across time and age, women had slightly higher levels of aggression than men. There was clear evidence for clustering within, and variation between families, both in overall aggression levels and in time effects.
Conclusions: We confirm earlier findings of a decline in aggression over the past decades. Not all individuals follow the downward trend over time for aggression to the same extent. Trends over time cluster within families, demonstrating that family factors are not only important to explain variation in aggression levels, but also in understanding differences between individuals in time trends.

Based on:
van der Laan, C. M., van de Weijer, S. G. A, Nivard, M. G., & Boomsma, D. I. (2021). Familial Clustering of Trends in Aggression. *Journal of Quantitative Criminology.* https://doi.org/10.1007/s10940-021-09523-8

FAMILIAL CLUSTERING OF TRENDS IN AGGRESSION

Aggression, behavior that intends to inflict harm on others (Berkowitz, 1993; Lorenz, 1966), can be encountered at home, at school, at the workplace, and elsewhere in the course of daily life. Aggression is a common type of human behavior (Tuvblad & Baker, 2011), that comes in various forms, ranging from non-physical, verbal, and relational aggression, to physical aggression (Crick & Grotpeter, 1995, Crick et al., 1999). Previous research has shown that males and females tend to differ in the form of aggression they portray. Physical aggression is more common in males, while indirect aggression seems to be more common in females (Hess & Hagen, 2006; Card, Sawalani, Stucky, & Little, 2008; Pickett et al., 2013; Beatton, Kidd & Machin, 2018; Thomson et al., 2019). Aggression typically decreases across the life-course (Alink et al., 2006; Cairns & Cairns 1989, 1994; Karriker-Jaffe et al., 2008; Loeber & Stouthamer-Loeber, 1998; Tremblay et al., 2004; Tremblay, 2010), while since the mid-nineties a declining trend in aggressive behavior at the population level has been observed (e.g., Pickett et al., 2013; Kann et al., 2016; Frøyland & von Soest, 2018). It is well known that individual differences in aggression tend to cluster within families (Margolin et al., 2016, Repetti et al., 2002; Besemer et al., 2017; van de Weijer et al., 2014), but a novel question is if in the change in aggressive behavior that is observed at the population level over time clusters within families. Here we address this question, after establishing that the general decline in aggression at the population level that is suggested in earlier studies replicates in Dutch society.

Trends in aggression over time

Most studies on trends in aggression over time have found declining levels over the last few decades. In the United States, the prevalence of adolescents having been in a physical fight decreased from 42.5% in 1991, to 22.6% in 2015 (Kann et al, 2016), while violent offending in adults decreased from 1991 to 2019 (US Department of Justice, 2019). In Norway, Frøyland and von Soest (2018) observed a decline in the prevalence of physical aggression in adolescence from 22.6% in 2007, to 12.8% in 2015. Pickett and colleagues (2013) found downward trends in physical fighting in adolescence from 2002 to 2010, in 19 out of 30 countries they investigated. Physical fighting in adolescents and violent crime in adults decreased in both males and females, but slightly faster in males, indicating that sex differences in physical aggression are decreasing (Beatton, Kidd & Machin, 2018). Clear statistics on trends in broader measures of aggression in everyday interactions (including non-physical, relational or verbal aggression) are lacking.

The downward trend we see in most physical aggression and violent crime statistics can be the consequence of multiple influences. Broadly, these can be divided into societal and developmental influences. Societal influences include period and cohort effects. Most studies suggest period effects are the main driving force behind the downward trends in aggression and violence (e.g. Fabio et al., 2006). Period effects affect the entire population, i.e., people of all ages and birth cohorts. Several specific period effects have been suggested in past studies that could explain such trends, for example increased policing and incarceration rates, lead poisoning, declining hard drug markets, technological advances, and economic circumstances (Farrel, Tilley & Tseloni, 2014).

In addition to potential period effects, the observed downward trend in aggressive behavior over the past decades could also be the consequence of cohort, or generational, effects. These are societal effects that affect only individuals from a certain birth-cohort. For example, Neugebauer, Hoek and Susser (1999) found an increased prevalence of antisocial personality disorder in a cohort of Dutch men prenatally exposed to severe maternal prenatal malnutrition during the harsh winter of 1944-1945 at the end of the Second World War. A sufficiently large birth cohort exposed to cohort-specific causes of aggressive behavior, could impact aggression trends on a populational level.

Trends in aggression over age

Temporal changes in behavior at the population level may also be due to developmental influences, defined as age, or life-course, effects. Physical aggression tends to peak at age 2 to 4 years and then declines to a lower level during the school years (Alink et al., 2006; Cairns & Cairns 1989, 1994; Karriker-Jaffe et al., 2008; Loeber & Stouthamer-Loeber, 1998; Tremblay et al, 2004; Tremblay, 2010). Social or relational aggression emerges in the preschool years, and continues through childhood and adolescence (Underwood, 2003). Because there are highs and lows in a nation's birthrate, the average age of a population will also change over time, and be reflected in aggression trends at the population level. For example, most western societies are ageing, having more older, and thereby less aggressive, individuals.

Familial clustering

Multiple studies have established that aggression, antisocial behavior, and violent crime cluster within families (Besemer et al., 2017; Margolin et al., 2016, Repetti et al., 2002; van de Weijer et al., 2014; Veroude et al., 2016;

Yu & Gamble, 2008). Such studies tend to rely on two approaches. One is to identify factors that vary between families, and assess their effects. This approach has the benefit that it provides information on specific family factors that are thought to play a role, which we then may target in interventions or policy. This approach has led to an extensive list of family factors related to psychosocial environment, social economic circumstances, and biological background, which can drive individual differences through a multitude of pathways (Tolan, Dodge, & Rutter, 2013; Farrington, Gaffney, & Ttofi, 2017; Labella & Masten, 2018). Many of these family influences are interrelated, and difficult to disentangle, adding to the complexity of the etiology of problem behavior. Guerra and Leidy (2008) provide an extensive overview of how individual characteristics, and contexts in which children are raised, can accumulate and interact, increasing or decreasing the risk of aggressive behavior. Dodge and Pettit (2003) illustrate the reciprocal nature of different influences on chronic conduct problems in a biopsychosocial model that includes a wide range of reciprocal influences situated in multiple domains: biological predisposition, parenting, peers, sociocultural context, and mental processes. When studying the effect of individual family factors, it is hard to account for their interdependency with other factors, and there is a risk of overestimating the causal effect of any single factor.

A second approach to study the role of family is to model family factors as one interrelated, but latent, construct. This approach does not require the identification and assessment of factors that vary between families, but does necessitate that outcome traits such as aggression are measured on multiple individuals from the same family. Using this design, family factors are defined as the combined total set of influences that lead to more similarities among individuals from within the same family than among individuals from different families.

Both approaches have their distinct advantages and disadvantages: the first approach requires measurement of the relevant family factors, while running the risk of not identifying important ones, but allows a design that does not require multiple individuals from the same family to take part in the study. The second approach does not require the identification of specific family factors, and allows us to get an understanding of the totality of influences that are specifically family based, without identifying these influences. If this second approach would offer little evidence for a significant 'between families' effect, the search to identify risk factors for aggression should be directed elsewhere.

Familial clustering of trends

Not only the levels or prevalence of violent and aggressive behavior is clustered within families, but developmental changes in these behaviors could also be more similar among family members than between families. Sampson and Laub, for example, argued that strong informal social bonds with family members lead to desistance from deviant behavior in adulthood (Sampson & Laub, 1994; Laub & Sampson, 2006). Persons from families with stronger social bonds between relatives might therefore all be more likely to desist, while those from families with weaker ties are more likely to persist in deviant behavior. In accordance with this line of reasoning, Van de Rakt et al. (2008) found intergenerational similarities in the developmental trajectories of criminal behavior among a large number of Dutch families and Eley, Lichtenstein and Moffit (2003) showed, in a twin study, that such similarities in developmental trajectories of aggressive antisocial behavior were largely mediated by genetic influences.

Although changes in criminal and aggressive behavior over the lifespan seem to cluster within families, this does not mean that time effects on aggression, reflecting changes in society, show similar clustering within families. We are not aware of any studies which investigated whether the beneficial downward trend in aggression over time that we see in the population also clusters within families. These trends in aggression over time may be explained by factors that affect all individuals in a population in a similar manner. In this case, individual differences in these time trends will be explained by the variation in initial aggression levels. However, time trends may also differ between individuals because changes in aggression over time reflect changes in underlying risk factors to which individuals are differently exposed, and which may be shared by family members. These changes can vary from economic circumstances, to changing role models or peers, to increased institutional control. If this is the case, time-trends are not a tide that raises or lowers all boats. Changes in society which may be beneficial to most, could harm others. Societal changes that affect families differently, may have a direct effect on aggression, or moderate the effect of other changes in society. A study by Odgers and colleagues (2015), for example, showed that living alongside more affluent neighbors predicts greater involvement in antisocial behavior among low-income boys. This suggests that non-conformity to beneficial economic change, may be one of the mechanisms that could lead to clustering of time effects on aggression within families. Regardless of the underlying mechanism, we expect time effects to differ between families. Consequently, we expect to see clustering of these effects within families.

The current study

In this study, we investigate the effect of time, age and gender on a broad behavioral measure of aggression, in a large cohort of Dutch multi-generation families with longitudinal data collected between 1991 and 2015, with a focus on assessing whether time and age effects cluster within families, i.e. are more similar within than between families. We use a novel approach in aggression research, by fitting a mixed effects model (Snijders & Bosker, 2012) to repeated measures of self-reported aggression of individuals within families, born between 1927 and 2001, aged 12-70 years. By specifying random effects in the model, the variance in overall aggression levels and in time and age effects is partitioned into within- and between family variance components. The random effects for families give an indication to what extent aggression levels and trends in aggression cluster within families. This informs us on the importance of family factors in explaining individual differences in trends in aggression over time and age.

METHODS

Data Collection & Participants

We analyzed longitudinal data, collected in twins, their parents and siblings who registered with the Netherlands Twin Register (NTR). Adult twins and their family members enter the Adult Netherlands Twin Register (ANTR) at different ages and newborn and young twins enter the Young Netherlands Twin Register (YNTR). When twins from the YNTR turn 18, they and their entire family automatically joins the ANTR and the ANTR data collection.

Survey data are collected in all participants every few years (Ligthart et al., 2019), in adolescent and adult participants by self-report, in children by parental and teacher report. From the ANTR databases, six waves of data collection were relevant for this study, as they contained the aggression scale from the Adult Self-Report (ASR, Achenbach & Rescorla, 2003). From these six waves of data collection, we selected all participants aged 12-70 years. The data were collected around 1991, 1995, 1997, 2000, 2009, and 2013. Twins and their siblings were asked to participate from 1991 onwards, and a selected group of their spouses from 2000. Parents of twins were invited to complete the ASR in 2009 and 2014. With each wave of data collection new families could enter the ANTR.

In YNTR twin pairs, three waves of data collection between 2004 and 2015 included aggression data from the Youth Self-Report (YSR; Achenbach & Rescorla, 2001). Twins were first approached when they were aged 12 to 18 and their siblings were invited to take part at the same time (see van Beijsterveldt et al., 2013). Thus, A family from the YNTR may have contributed YSR aggression data for the first time in 2006 from 16-year old twins and an additional sibling, and in 2009 through the ANTR survey 5, when the twins were 19-years old. Table 3.1 presents an overview of the surveys and sampling description (see also; Boomsma et al., 2002, 2006; Ligthart et al., 2019). Research ethics committee approval was received for each survey.

Table 3.1. Sample description: Number of participants at each wave of data collection in the Adult and the Young Netherlands Twin Register, and the number of measures of these same individuals on subsequent waves of data collection

Adult Self Report (ASR) in ANTR							
	Observations		**N observations on following surveys**				
Survey (year)	**(total)**	**Age range (M)**	**ASR 2**	**ASR 3**	**ASR 4**	**ASR 5**	**ASR 6**
ASR 1 (1991-1993)	3,325	12-24 (17.21)	1,870	1,253	1,227	973	822
ASR 2 (1995-1996)	3,326	14-28 (19.30)		1,807	1,767	1,259	1,042
ASR 3 (1997-1999)	4,685	12-70 (26.10)			2,875	1,981	1,532
ASR 4 (2000-2003)	6,655	12-70 (29.70)				3,128	2,456
ASR 5 (2009-2013)	14,639	14-70 (40.15)					7,553
ASR 6 (2013-2015)	16,143	15-70 (39.61)					
Youth Self Report (YSR) in YNTR							
	Observations		**N observations on following surveys**				
Survey (year)	**(total)**	**Age range (M)**	**YSR 2**	**YSR 3**	**ASR 5**	**ASR 6**	
YSR 1 2005-2013	11,108	12-26 (14.59)	4,204	0	929	2,533	
YSR 2 2005-2014	8,073	12-28 (16.74)		410	1,479	2,936	
YSR 3 2004-2008	1,511	12-35 (18.31)			745	623	

Abbreviations: ASR = Adult Self Report, YSR = Youth Self Report.
Note: When YSR twin participants turn 18, they, and their families, can take part in the adult (ASR) surveys.

Our dataset consisted of 69,465 measures from 40,400 participants, from 15,437 families (see Table 3.2 and 3.3). Around 52 percent of the participants come from either monozygotic twin pairs (9,228; 22.8%) or dizygotic pairs (11,760; 29.1%). The remainder of the participants are twins from incomplete pairs, siblings, parents, spouses or offspring of twins. Approximately 43 percent of all participants completed more than one survey, 18 percent

completed more than two surveys. Because not all NTR participants were invited for all surveys, as described above, the composition of the sample differs over time.

Table 3.2. Description of included families: Number of family members, observations and generations

N	N	N	Generations in family	
family members	families	observations	1	2
1	3,628	4,291	3,628	-
2	5,115	14,478	4,657	458
3	2,874	14,478	1,632	1,242
4	2,068	15,671	409	1,658
5	1,211	12,551	114	1,097
6 - 16	541	7,996	67	473

Note: One generation families consist of family members of the same generation (i.e. siblings and twins), two generation families consist of two generations (i.e. parent(s) and offspring). Additionally, two three-generation families were included in the analyses (not shown in Table 3.2), consisting of grandparent(s), parent(s), and offspring.

Table 3.3. Sample characteristics

				IRT Aggression	
Group	N subjects	N measures	Mean age	Mean	SD
Total	40,400	69,465	28.70	0.00	0.82
Male	16,588	26,907	28.60	-0.05	0.81
Female	23,806	42,548	28.76	0.03	0.82

Note: For six participants gender was missing, or the participant changed gender. Mean age is calculated over all measures.

Measurements

All participants completed Achenbach System of Empirically Based Assessment self-report questionnaires (ASEBA; Achenbach et al., 2017), either the Youth Self-Report (YSR; Achenbach & Rescorla, 2001) or the Adult Self-Report (ASR; Achenbach & Rescorla, 2003). In the early ANTR surveys, the YASR (Young Adult Self Report) was administered.

A score based on nine overlapping items that were present in all surveys was analyzed (see Table 3.4). All items were scored on a three-level scale: 0= never, 1= sometimes, 2= often. Aggression scores were defined by Item-Response Theory (IRT; Embretson & Reise, 2000) and calculated with the

Generalized Partial Credit Model (GPCM) in *R*, with the *mirt* package (2012). GPCM is an Item Response Theory model, developed to analyze polytomous data. This IRT-aggression score has benefits over a simple sum-score, because it appropriately weights the relative contributions of individual items to a scale with a more favorable distribution.

Table 3.4. Aggression items: response rate and IRT weight

	Response					
Items	***Never***	***Sometimes***	***Often***	**Mean ♂**	**Mean ♀**	**IRT weight**
Argues	44,389	23,466	1,493	0.33	0.41	0.641
Mean to others	56,688	12,060	468	0.20	0.18	0.583
Fights	66,828	2,165	215	0.06	0.02	0.678
Attacks	67,536	1,452	73	0.04	0.02	0.760
Screams	60,465	8,042	800	0.11	0.16	0.714
Stubborn	34,277	30,296	4,731	0.55	0.59	0.679
Mood changes	43,897	21,753	3,614	0.32	0.48	0.610
Temper	55,772	12,227	1,382	0.23	0.21	0.779
Threatens	67,646	1,247	166	0.04	0.01	0.663

Note: Observations with more than 2 missing items were removed from the analyses. IRT weights can be interpreted as factor loadings.

Analyses

We fit a three-level mixed effects model (e.g. Snijders and Bosker, 2012) to the longitudinal IRT data, with the package *nlme* (Pinheiro et al., 2017) in *R* (version 3.4.2; R Core Team, 2017). We opted for a multilevel modeling approach because we are interested in the effects of age and time on aggression, and the within and between family variance of these trends. By modeling the within and between family variance we get an estimate of the extent to which family characteristics play a role in individual variance in aggression trends. The family characteristics are in this case a latent (i.e. not observed) construct that embodies all factors that lead to more similarities within- than between families. The mixed modeling approach allows for this, while also considering dependency between repeated measures.

Our model accounts for three levels of variance that are present in the data, namely variation between repeated measures, variation between individuals, and variation between families. The model is called a mixed effects model because both fixed and random effects are estimated. Fixed effects are population level effects that apply to each measure. Random

effects are effects that allow for variation of effects at group level. Minica˘ et al. (2015) showed that a similar approach generated correct standard errors in a design with varying levels of family resemblance. In both models we use *z*-scores of the continuous predictors (age and time). Gender is included as a fixed effect to assess gender differences in overall aggression levels and time and age trends. An autoregressive correlation structure is applied in the model at the individual level to account for dependence between measures. The autoregressive correlation structure assumes that measures in closer temporal proximity are more strongly related than more distant measures. A correlation between the random effects is calculated, giving an indication of the association between the estimated random intercept and random slope for time. Explained variance, R^2, is calculated with the package *MuMIn* (Barton, 2018) in *R*.

The outcome (aggression) is predicted by an intercept that can vary for each individual in each family, slopes for time and age that can vary for each family, a slope for gender, and non-linear slopes for age (age^2) and time ($time^2$). Parameter estimates with random effects are underscored:

Level 1:

$$AGG_{tij} = \underline{\beta}_{0ij} + \underline{\beta}_{1j} Age_{1tij} + \beta_2 Age^2_{1tij} + \underline{\beta}_{3j} Time_{2tij} + \beta_4 Time^2_{2tij} + \beta_5 Sex_{3tij} + first\ order\ interactions + \underline{\varepsilon}_{tij}$$

In this notation, *AGG*tij is the aggression outcome for measurement *t* in individual *i* in family *j*, $\underline{\beta}_{0ij}$ is the intercept for each individual *i* in family *j*, β_{1j} is the linear slope for age in family j, β_2 is the non-linear slope for age, β_{3j} is the linear slope for time in family *j*, β_4 is the non-linear slope for time, β_5 is the slope for gender, and $\underline{\varepsilon}_{tij}$ is an error term for each measurement *t*, for individual *i*, in family *j*. All first order interactions are included in the model. The random effect for the intercept for each individual is estimated by the global intercept and a random effect for individual *i*, and makes up Level 2 of our model:

Level 2:

$$\underline{\beta}_{0i} = \beta_0 + \underline{u}_{0i}$$

This random term accounts for dependency of repeated measures within individuals. The random effect for the intercept for each family is estimated by the individual intercept and a random variable for family *j*. The random

effects for the slopes of time and age for each family are estimated by the global slopes for time and age, and random effects for family *j*. These family-level effects form Level 3 of our model:

Level 3 intercept: $\underline{\beta}_{0ij} = \underline{\beta}_{0i} + \underline{u}_{0j}$

Level 3 slopes:

Slope for age: $\underline{\beta}_{1j} = \beta_1 + \underline{u}_{1j}$

Slope for time: $\underline{\beta}_{2j} = \beta_2 + \underline{u}_{2j}$

The random family effects, for the intercept and the slope of time and age, are the focus of our study. By specifying random effects in our model, variance is partitioned into within- and between family variance components. The model estimates of the random effects for families are a measure of between family variance. This informs us on how much individuals from different families differ in their intercept and slopes for time and age, compared to family members. Thus, the random intercept for families gives an indication to what extent aggression levels cluster within families. Similarly, the random effects for the slopes of time and age give an indication to what extent these effects cluster within families. No discrimination is made between different family relations when estimating random family effects. Thus, the family effects reflect overall resemblance between family members, including partners, siblings, parent–offspring or twins. The correlation between the random parameter estimates gives an indication of how aggression scores (high or low) are related to the increase or decrease of these scores over time and age. We tested the reliability of our model in estimating variation at the family level in both intercept and slopes with simulated data. The model was effective in estimating variation between families. Type I error rates for the random intercept and random slopes were investigated in a subset of 420 individuals from 100 families. Simulations were run 1000 times, both with a normally distributed outcome and with the empirical IRT distribution in the outcome. Type I error rates were similar for both outcomes, but slightly lower than expected, ranging from 1.2% to 1.5%.

The effect of time reflects both period and cohort effects. To get an idea of the relative importance of period and cohort effects, we investigate linear trends within four age categories: 12 to 25 years, 26 to 40 years, 41 to 55 years, and 56 to 70 years. These four age ranges represent four birth cohorts

in our data. If change in aggression in our data would be due to changes in specific birth cohorts, we would see clear differences in aggression trends over time between the four cohorts. With this simple method we aim to establish whether change is cohort-specific or due to period effects. We fitted the linear relation between time and aggression within the four groups with the *lm* method in *ggplot2* (Wickham, 2009).

RESULTS

Our dataset consisted of 69,465 aggression measures collected between 1991 and 2015 from 40,400 participants (58% female), who were part of 15,437 families (see Tables 3.2 and 3.3). On average, aggression declined between 1991 and 2015, by 0.87 *SD* (Figure 3.1). The downward trend was found for both sexes, with women scoring on average 0.08 *SD* higher than men, which can be explained by higher scores on items related to non-physical aggression (Table 3.4). Figure 3.2 shows that aggression decreased over the beginning of the life course, from age 12 to age 34, by 0.94 *SD*. From age 34 to age 70, only a slight increase in aggression, of 0.06 *SD,* was visible in males. We are reluctant to interpret this slight increase, as this could be due to overfitting because of smaller sample size at old ages.

Because the model does not allow for simultaneous analysis of time and cohort effects, the results did not inform us on which of these effects drove the effect of time. To get an idea of the relative importance of period and cohort effects, we plotted the linear effect of time on aggression within four age categories; 12 to 25 years, 26 to 40 years, 41 to 55 years, and 56 to 70 years (Figure 3.3). A clear downward trend was visible within each age category. This suggested that period effects predominantly drove the downward trend, across ages and birth-cohorts.

The fixed and random effects together explained 52 percent of the variance in aggression, with the fixed effects alone explaining 14 percent of the variance in aggression, i.e. random effects accounted for the largest part of the variance in aggression measures. The random effects included a random intercept for individuals to account for dependence between repeated measures, and a random intercept and random slopes for time and age to investigate variation between, and clustering within families. Table 3.5 shows the estimates for the fixed and random effects.

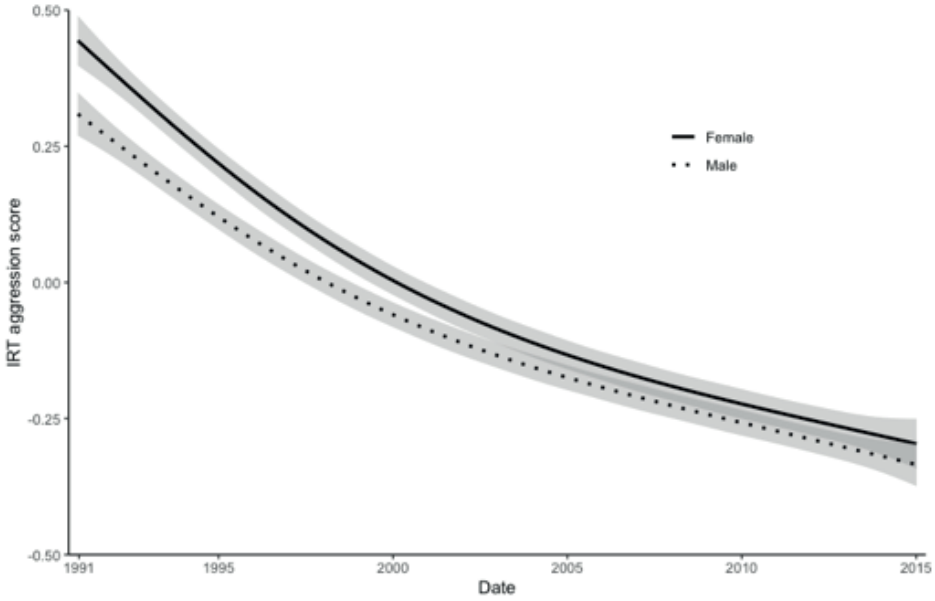


Figure 3.1. Predicted effect of time: each time point aggregates predicted aggression scores over ages and birth cohorts

Note: The grey banners depict the 95% Confidence intervals.

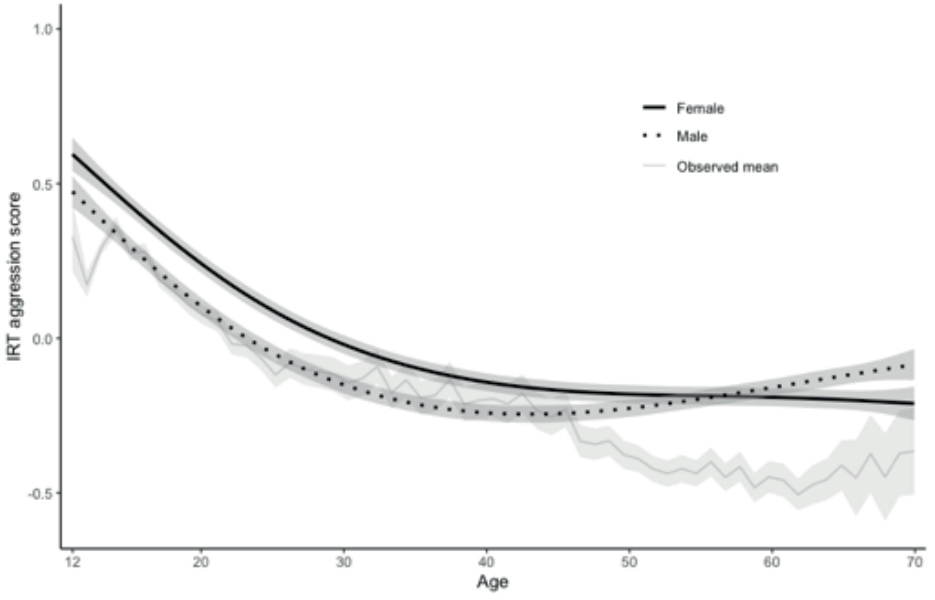


Figure 3.2. Predicted and observed effect of age on aggression

Note: The grey banners depict the 95% Confidence intervals.

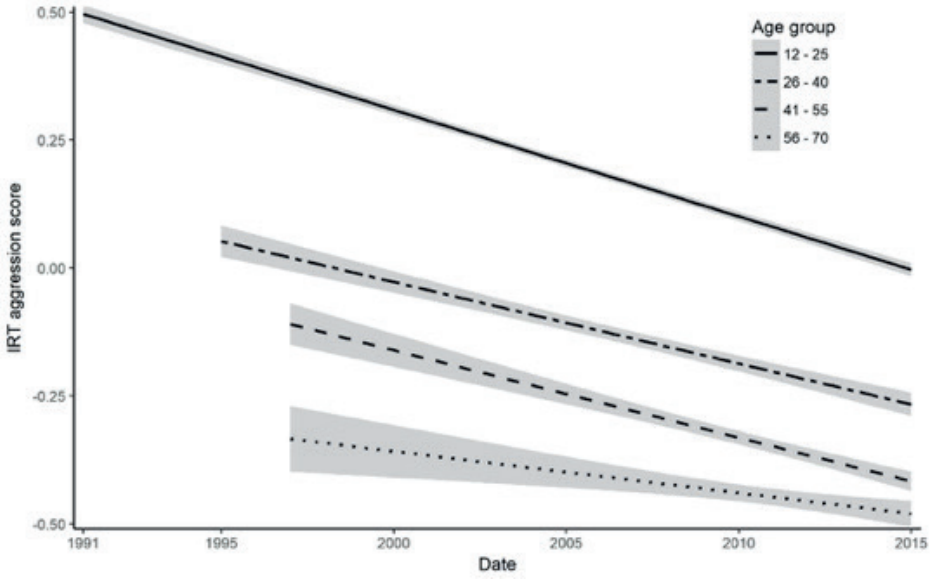


Figure 3.3. Linear effect of time on aggression within 4 age categories; 12 to 25 years, 26 to 40 years, 41 to 55 years, and 56 to 70 years

Note: The grey banners depict the 95% Confidence intervals.

By including random effects for families, variance was partitioned into within and between family variance. The random intercept for families is a measure of between family variance: $SD = 0.302$, $\Delta\chi^2(3) = 1333.739$, $p < 0.001$. This indicated that variance in aggression levels between individuals from different families was larger than the variance between individuals within the same family, or stated otherwise, that aggression levels clustered within families.

Our main interest is in the variance of the effects of time and age, which were also partitioned into within- and between family variance components. Between family variance of the effect of time on aggression was estimated by a random slope for time: $SD = 0.116$, $\Delta\chi^2(3) = 96.825$, $p < 0.001$, indicating that variance in the effect of time on aggression between individuals from different families was larger than the variance between individuals within the same family. Thus, members of the same family resembled each other in the magnitude and direction of the effect of time. Moreover, the model with a random slope for time for families fit significantly better than the same model with a random slope for time for individuals, $\Delta\chi^2(1) = 30.592$, $p < 0.001$.

The estimate of the random effect of time was not related to the estimate of the random intercept for families, $r = .039$, meaning that these estimates are interpretable separately.

Between family variance of the effect of age on aggression was estimated by a random slope for age: $SD = 0.102$, $\Delta\chi^2(3) = 416.744$, $p < 0.001$, indicating family resemblance in age effects. We could not compare the fit of the model with a random age effect for family with a model with a random age effect for individuals, as the model with a random age effect for individuals did not converge. Contrary to the random effect of time, the estimate of the random effect of age was negatively associated with the estimate of the random intercept for families: $r = -.912$, making an interpretation of the random effect of age separate from the random intercept impossible. This is a result of a strong relationship between age and aggression. Families with higher aggression scores are more likely to have a stronger decrease over age.

DISCUSSION

We investigated whether trends in aggression cluster within families, making this the first study that combines the analyses of trends and family clustering of aggressive behavior. We investigated whether the effects of time and age on aggression cluster within families, after first establishing that a downward trend in aggressive behavior is replicated in the Netherlands between 1991 and 2015.

Levels of self-reported aggression were calculated by IRT (Embretson & Reise, 2000). Overall, IRT aggression scores decreased by 0.87 *SD* from 1991 to 2015. Our model did not allow for simultaneous estimates of period and cohort effects, but aggression also decreased within separate age categories, suggesting that mainly period effects play an important role: changes in society that affects the entire population across ages and birth-cohorts. If cohort effects were more important, we would not expect a similar decrease within different age categories. Identifying which specific period effects drive the trend we see is not an easy feat. It is likely that a large number of diverse factors play a role. Farrell, Tilley and Tseloni (2014) reviewed 17 hypotheses to identify the cause for the observed drop in criminal and violent behavior. They included hypotheses related to increased policing and incarceration rates, lead poisoning, declining hard drug markets, technological advances, increased security and economic

Table 3.5. Parameter estimates and explained variance for the fixed and random effects of time, age, and gender on aggression

Random effects	***SD***		***Interpretation***
Family			
Intercept	*0.302*	***	Between family variance > within family variance. Aggression differs between, and clusters within, families
Time	0.116	***	Between family variance > within family variance. The effect of time differs between, and clusters within, families
Age	0.102	***	Between family variance > within family variance. The effect of age differs between, and clusters within, families
Individuals within families			
Intercept	0.381	***	Between individual variance > within individual variance. Dependency of repeated measures. Repeated measures in an individual are more alike than measures in different family members.
Residual	*0.570*		*Variance not accounted for by the random effects.*
Fixed Effects	***Estimate***		***95% CI***
Main effects			
Intercept	-0.157	***	[-0.181, -0.134]
Time	-0.066	***	[-0.084, -0.047]
Age	-0.320	***	[-0.340, -0.300]
$Gender_a$	0.107	***	[0.084, 0. 130]
$Time^2$	-0.015		[-0.030, 0.001]
Age^2	0.168	***	[0.148, 0.188]
First order interactions			
Time * Age	0.141	***	[0.126, 0.156]
Time * Age^2	-0.085	***	[-0.096, -0.073]
Time * $Time^2$	-0.012	**	[-0.020, -0.004]
Age * $Time^2$	0.069	***	[0.056, 0.083]
Age * Age^2	-0.032	***	[-0.040, -0.024]
Age^2 * $Time^2$	-0.012	*	[-0.021, -0.002]
Time * $Gender_a$	-0.012		[-0.028, 0.004]
Age * $Gender_a$	-0.022	*	[-0.041, -0.003]
$Gender_a$ * $Time^2$	0.012	***	[-0.001, -0.024]
$Gender_a$ * Age^2	-0.028	***	[-0.042, -0.013]
	R^2		
Fixed effects	0.14		
Fixed & random effects	0.52		

Note: Time and age were standardized for the analyses. Due to the interactions between the fixed effects, the fixed effects are not interpretable separately. For the predicted scores based on the fixed effects see Figure 3.1 and 3.2.
* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

$_a$ Gender was coded male= 0, female= 1.

circumstances, but did not find much proof to either debunk or corroborate the hypotheses. It was beyond the scope of the current study to directly test specific factors that may drive the observed time effects, but this remains an important topic for criminological research.

In the mixed-effects model, variance in aggression levels was partitioned into within- and between family variance. We found that aggression levels cluster within families: family members are more alike in aggressive behavior than individuals from different families. This is in line with multiple studies on aggressive, criminal and antisocial behavior (Besemer et al., 2017; Margolin et al., 2016; Repetti et al., 2002; van de Weijer et al., 2014; Veroude et al., 2016). Factors shared by family members thus may help in explaining individual differences in aggression, and previous research has identified several of those factors that likely play a role in driving individual differences in aggression, including genetic risk, low income, harsh parenting, and family violence (Margolin et al., 2016; Repetti et al., 2002; Labella & Masten, 2018). We added a new element to family analyses, namely the analyses of trends and family clustering. We investigated whether the effects of time and age on aggression also cluster within families. Variance in slopes was partitioned into within- and between family variance. We found that for the effect of time, family members are more alike in trends over time compared to unrelated individuals. So, even though at the population level, average scores are declining, not everyone in society is benefitting from this trend. There are two ways to interpret these results. Risk factors shared by family members could change over time, leading to a similar effect on family members. Alternatively, societal effects may be moderated by family factors, leading to different outcomes in different families. Under both scenarios and interpretations, family factors are not only important in understanding differences in aggression levels, but also in understanding change over time. The overall downward trend is, thus, not necessarily beneficial for everyone. The familial resemblance in how far individuals diverge from the overall trend, indicates that certain families or groups are differentially affected by changes in society. One possible mechanism could be variation in the benefits of economic progress. Not only is poverty a well-established risk factor for aggression, but also differences in 'relative poverty": a poor family, living in close proximity to a more wealthy family, has a further increased risk (Odgers et al., 2015; Nieuwenhuis et al., 2017). Thus, if other families in the neighborhood benefit from economic progress, the effect of poverty on aggression may increase.

Regardless of the mechanisms behind the clustering of trends, our study provides strong indications that even with beneficial changes in society, we should not lose focus on families when trying to understand or change aggressive behavior. Families are important in understanding the mechanisms that drive aggressive behavior now, and in the future. Because the estimate of the random slope for time was not related to the estimate of the random intercept, factors that lead to clustering of aggression levels within families may not be the same factors that lead to clustering of time trends. Risk factors for aggression that reside in the family system, may not be static but dynamic, changing over time as society as a whole changes. Policymakers should be aware that theories on aggression might not translate over time, and that the importance of specific factors may change as society changes.

The age-aggression curve observed in the present study resembles previous results, with a decrease in aggression over age that gradually levels off from early adulthood onwards (Alink et al., 2006; Cairns & Cairns 1989, 1994; Tremblay, 2010; Tremblay et al., 2004). We also found that variance between non-related individuals was larger than variance between family members for the effect of age, indicating clustering of age effects within families. Liu, Lewis and Evans (2013) discuss risk factors for aggression pertaining to several developmental stages. They conclude that different risk factors, of which several may be shared within families, have unique influences at specific ages. Thus, familial risk factors not only contribute to overall levels of aggression, but also to the development of aggression over age. In line with this finding, Eley, Lichtenstein and Moffitt (2003) also showed that continuity of aggressive and nonaggressive antisocial behavior in childhood was mediated by genetic and shared environmental factors. Our results indicated that the estimate of the random slope for age, reflecting clustering of age effects in families, was dependent on the estimate of the random intercept, reflecting clustering of aggression levels in families. This indicates that we cannot interpret the clustering of age effects separately from the overall aggression levels. In other words, high levels of aggression are related to a large age effect, presumably because starting levels are higher, and lower levels are related to a smaller age effect.

Results further indicate slightly higher overall aggression in females than in males. A close inspection of the aggression scores in males and females revealed that this difference was mainly because females scored higher on items related to non-physical or relational aggression. Previous research

is inconsistent on gender differences in indirect aggression. For example, Crick and Grotpeter (1995) found that relational aggression was more prevalent among adolescent girls than boys, and Thomson and colleagues found similar higher levels of indirect aggression in adult females compared to males. However, in a large meta-analysis on 148 studies on child and adolescent aggression, Card, Sawalani, Stucky, and Little (2008) found trivial gender differences in what they call 'indirect' aggression. Several factors may cause gender differences in aggression, including biological factors, for example differences in physical strength (Björkqvist, 1994) and hormonal differences (Björkqvist, 2018); and socio-cultural factors, for example different role models (Underwood et al., 2008), differences in the number and closeness of social relationships (Maccoby, 1990) and differences in the social acceptance of aggression (Underwood, 2003). We did not find clear gender differences in trends in aggression over time and age, indicating that the factors driving the overall downward trend we see over time, affects males and females similarly.

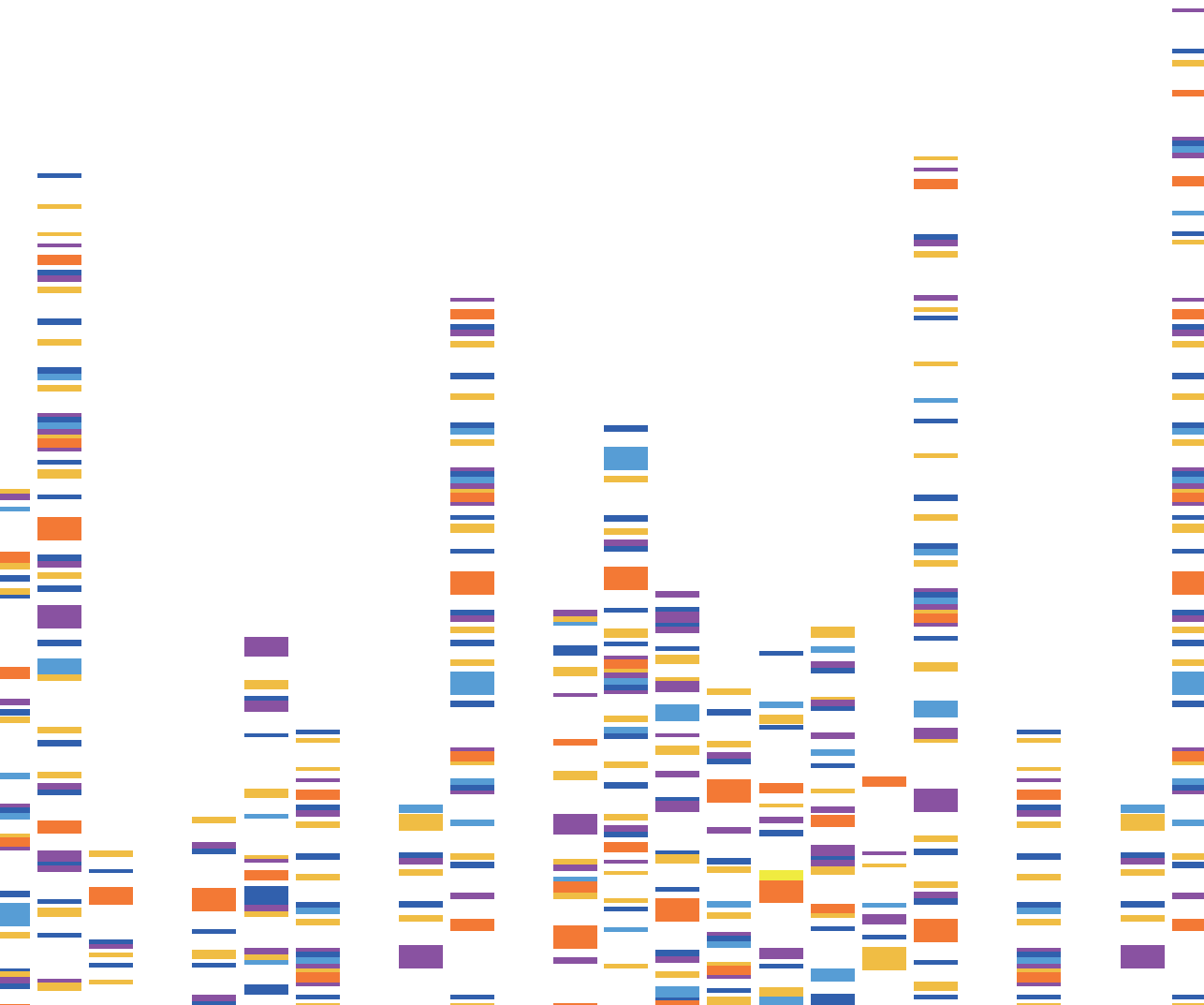
Although our study provides important new insights on the development of aggressive behavior over time, it also has several limitations. Our longitudinal study design is inevitably accompanied by drop-out, both 'by design', because not all participants were asked to take part in all surveys, and due to nonresponse. Nonresponse may partly explain the negative aggression trend that we found, if more aggressive respondents are more likely to drop out than less aggressive respondents. Additional analyses showed the effect of drop-out to be small, if not absent. Participants with multiple measures scored slightly lower on aggression in most age categories, however, our study allows for new inflow of participants at each wave, so that the effects of non-response are not limited to later surveys. Analyses with only the first measure of each participant yielded similar results to those reported, indicating that the downward trend is not driven by repeated measures of less aggressive individuals.

A second limitation is that the adult self-report (ASR) surveys slightly differ from the youth self-report (YSR) surveys. To ensure comparability of measures, we included the nine aggression items that are the same in both surveys.

Finally, family environments may increase the likelihood of displaying aggressive behavior, because of close proximity and more interaction between individuals compared to people living alone. Since we analyzed

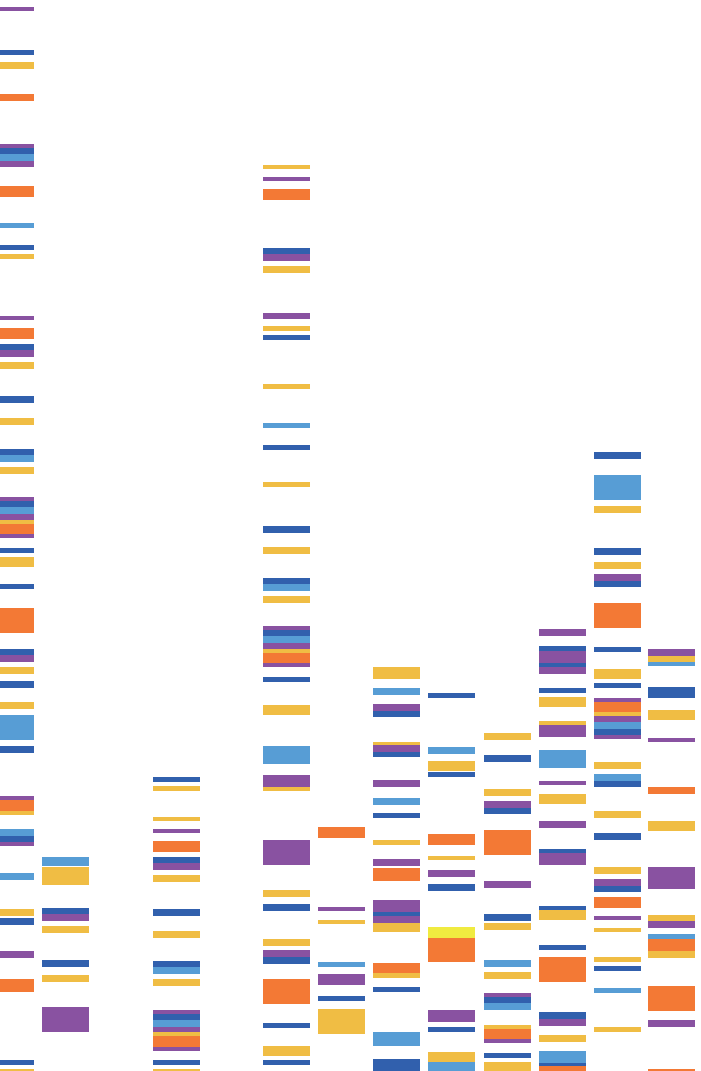
family data, this may have slightly biased the results. To assess this possible bias, we tested for differences in aggression levels between twins and their non-twin siblings, and found no significant differences within 10-year age groups. Moreover, twin pairs may interact more closely with each other than siblings. However, there is no evidence for an increase in similarity between twins compared to non-twin siblings when using self-report measures (Plomin & Daniels, 2011; Mark, Pike, Latham & Oliver, 2017). We also reran the analyses without monozygotic twins, and found similar results, indicating that the relatively small number of monozygotic twin pairs did not drive the effects we see.

In summary, we replicated that aggression decreases in the period between 1991 and 2015, with an average decline in aggression scores of 0.87 SD. The decrease over time was present across all ages in the sample. Aggression clustered within families. Thus, factors shared by family members can help explain individual differences in aggression. The effect of time on aggression also cluster within families. This indicates that not all individuals follow the beneficial trend we see in the population, and demonstrates that family factors are important in understanding and explaining differences between individuals in their aggression development over time.



4

Continuity of Genetic Risk for Aggressive Behavior Across the Life-Course



ABSTRACT

We test whether genetic influences that explain individual differences in aggression in early life also explain individual differences across the life-course. In two cohorts from The Netherlands (*N*= 13,471) and Australia (*N*= 5,628), Polygenic Scores (PGSs) were computed based on a genome-wide meta-analysis of childhood/adolescence aggression. In a novel analytic approach, we ran a mixed effects model for each age (Netherlands: 12-70 years, Australia: 16-73 years), with observations at the focus age weighted as 1, and decaying weights for ages further away. We call this approach a 'rolling weights' model. In The Netherlands, the estimated effect of the PGS was relatively similar from age 12 to age 41, and decreased from age 41 to 70. In Australia, there was a peak in the effect of the PGS around age 40 years. These results are a first indication from a molecular genetics perspective that genetic influences on aggressive behavior that are expressed in childhood continue to play a role later in life.

Based on:
van der Laan, C.M., Morosoli-Garcia, J.J., van de Weijer, S.G.A., Colodro-Conde, L., ACTION consortium, Lupton, M.K., ... Boomsma, D.I. (2021). Continuity of genetic risk for aggressive behavior across the life-course. *Behavior Genetics, 51*, 592-606. https://doi.org/10.1007/s10519-021-10076-6

CONTINUITY OF GENETIC RISK FOR AGGRESSIVE BEHAVIOR ACROSS THE LIFE-COURSE

Aggression is broadly defined as common human behavior that intends to cause harm, by verbal, psychological, and physical means, to others (Berkowitz, 1993; Baron & Richardson, 1994; Lorenz, 1966). Physical aggression tends to peak at age 2-4 years and then decreases (Alink et al., 2006; Cairns & Cairns 1989, 1994; Karriker-Jaffe et al., 2008; Loeber & Stouthamer-Loeber, 1998; Tremblay et al, 2004; Tremblay, 2010), as neurological, cognitive and social development empower children with other means to get what they want. Social or relational aggression emerges in the preschool years, continues through childhood and adolescence and subsequently declines in adulthood (e.g. Underwood, 2003).

The relative positions in terms of aggression (i.e. rank order) in the population persist across the life-course (Pulkkinen & Pitkänen, 1993; Tuvblad & Baker, 2011). In other words, the most aggressive child often grows up to be the most aggressive adult (Farrington, 1989). There has been some debate about the continuation of individual differences in aggression from childhood to adulthood. Moffit (1993) argued that this statistical continuation is driven by a small number of highly aggressive individuals in a population who remain aggressive throughout their lives, the 'life-course persistent' individuals. The rest, she argues, are the 'adolescent limited' type, for whom aggressive behavior is limited to adolescence. Although it is clear that the 'life-course persistent' individuals explain part of the stability in aggression, Huesmann, Dubow, and Boxer (2009) showed that most individuals retain their relative position in a population, regardless of their starting position. Several factors have been identified that help explain individual differences in continuity of aggression, such as parenting, peers, socioeconomic and cultural context, mental processes and genetic predisposition (Boomsma, 2015; Farrington et al., 1989; Labella & Masten, 2018; Murray & Farrington 2010; Tolan, Dodge, & Rutter, 2013; Vuoksimaa et al., 2020).

Twin and family studies, mostly focusing on children, indicate that genetic factors explain around 50 percent of the variation in aggression (Veroude et al., 2016). Across the lifespan, heritability estimates of aggression and antisocial behavior seem to increase somewhat from childhood through adulthood, as the importance of shared environmental effects decreases (Tuvblad and Baker, 2011; Waltes et al., 2016, Odintsova et al., 2019). Although individuals retain their genetic make-up throughout their lives,

this does not necessarily imply that the same genetic variants play a role in aggression across the life-course. Studies with longitudinal twin designs show that genetic factors contribute significantly to the stability of aggression during preschool age, school age, and puberty (van Bijsterveldt et al., 2003; Porsch et al., 2016). These results led us to test the hypothesis that genetic variants that are expressed on aggression during childhood and adolescence also are significantly associated with aggression later in the life-course.

Odintsova and colleagues (2019) published an extensive overview of the current state of genomics aggression research, concluding that clear genome wide significant effects have not yet been found in genetic association studies (GWAS). This is partly attributable to the fact that aggression, like many other complex human behaviors, is influenced by a multitude of individual genetic variants, each of which likely has a small effect. From this 'polygenic' genetic architecture, arises the need for very large GWAS sample sizes. Ip and colleagues (2021) conducted a genome wide meta-analysis (GWAMA) of aggression phenotypes in children and adolescents, aged 3 to 18 years. In a GWAMA, results from GWAS in multiple cohorts are combined with the aim to increase statistical power to find associations between a genetic marker (usually a single nucleotide polymorphism, i.e. SNP) and an outcome (phenotype). If phenotypes and genetic effects are comparable across cohorts from different ages and backgrounds, small effects of SNPs that do not attain significance in a single cohort may be genome wide significant in the GWAMA. In the Ip et al. paper (2021), a total of 29 cohorts contributed 163 univariate GWAS to the early life aggression GWAMA. This resulted in a total of 328,935 observations from 87,485 unique individuals, aged 3 to 18 years. Observations were across multiple raters, coming from teachers, parents, and self-reports. The Ip et al. (2021) aggression GWAMA is the largest childhood aggression GWAS to date, but no single genome-wide significant hits were observed. Despite this lack of single significant hits, Ip et al. (2021) demonstrated that polygenic scores (PGSs), which sum the effects of a range of genetic markers, with markers included based on whether their *p*-value from the GWAS clears any of 16 thresholds between P = 1 and P < 1.0E-5, explained between 0.036 and 0.44 percent of the phenotypic variance in aggression in a hold-out sample of 7 year-old Dutch children (*N*=4,491). In an Australian hold-out sample, childhood PGSs explained up to 0.2 percent of retrospectively assessed childhood conduct disorder. PGSs performed best when markers where included with relatively lenient p-value thresholds, indicating the polygenic nature of aggression

phenotypes. Although effect sizes were small, we expect the effect to be large enough to test the hypothesis that there are continuing genetic effects across the life-course.

The current study

In this study, we test the hypothesis that genetic risk factors, measured as DNA variants associated with increased aggressive behavior in early life (Ip et al., 2021), increase the risk of aggression across the life-course. We quantified the contribution of a large number of variants by computing PGSs, and tested their association with aggression in two cohorts from two different countries, namely The Netherlands and Australia. We introduce a novel method to assess differences in genetic influences across the life-course. In this approach we assess the effect of a PGS on aggression at ages 12 to 70 in The Netherlands, and 16 to 73 in Australia, by specifying 'rolling weights' for age. Within the framework of a linear mixed model, we model the effect of the PGS at each age represented in the data. At each age, we include phenotype information from surrounding ages. The phenotype information is weighted, where weights are centered at the focus age and decay further away from that center. With this method, sample size differences between ages are small, because more information than just the focus age is taken into account. Thereby, we mitigate the risk that sample size differences between ages drive the effects we find. If, for example, there are only few observations at age 25, we can still use information on adjacent ages to imply the effects at age 25. A significant contribution in adults, of PGSs that are based on a discovery study in children and adolescents, would suggest a partially heritable origin of the stability of individual differences in aggression.

METHODS

Participants

Dutch participants are registered with The Netherlands Twin Register (NTR; Ligthart et al., 2019). Phenotype data on aggression were collected by survey in six of the data collection waves (1991, 1995, 1997, 2000, 2009, 2014) for twin families who registered as part of the Adult Netherlands Twin Register (ANTR). Twins, whose parents registered them as newborns as part of the Young Netherlands Twin Register (YNTR), and their siblings provided self-ratings of aggression at various ages of the twins. The total phenotyped sample in which aggression scores were computed (Table 4.1), consisted of

families with twins, their siblings, spouses, and parents. In the final analyses only the genotyped individuals could be included. This genotyped sample included 29,454 measures from 13,471 genotyped individuals (62% female) aged 12 to 70 years (Table 4.2, Figure 4.1). In total, 8,705 of these individuals completed more than one questionnaire. All genotyped individuals were from European ancestry, identified based on the top ten 1000-Genomes PCs (Abdellaoui et al., 2013).

Table 4.1. The Netherlands Twin Register: Collection of aggression data in adolescent and adult twins, their sibs, spouses and parents (1991-2014) and in young twins and their siblings (2005-2014)

Year	Observations	Mean age (SD)
1991	3,325	17.95 (2.24)
1995	3,342	19.98 (3.10)
1997	4,714	26.73 (10.46)
2000	6,684	30.48 (10.75)
2009	14,798	41.44 (15.40)
2014	16,092	40.16 (14.61)
2005-2014 (Age twins: 14)	11,080	15.35 (1.54)
2005-2014 (Age twins: 16)	8,075	17.43 (1.60)
2005-2008 (Age twins: 18)	1,516	18.88 (1.94)

Note: This table includes all phenotype observations that were included to calculate IRT aggression scores.

Table 4.2. The Netherlands Twin Register: Characteristics of genotyped/phenotyped sample

				IRT Aggression	
	***N* subjects**	***N* measures**	**Mean age (*SD*)**	**Mean**	***SD***
Total	13,471	29,454	31.35 (15.33)	-0.02	0.85
Male	5,062	10,426	31.34 (16.07)	-0.07	0.84
Female	8,409	19,028	31.35 (14.91)	0.01	0.85

Note: IRT Aggression= item response theory aggression score, SD= standard deviation.

Australian data came from studies on health and wellbeing collected at QIMR Berghofer Medical Research Institute (QIMRB). A total of 5,628 genotyped participants from 2,983 families from the Brisbane Longitudinal Twin Study (Wright and Martin, 2004), Young and Well (16UP study, Mitchell et al., 2021) and Twenty Five and Up (25UP study, Mitchell et al., 2019) Genetics and Human Agency study (GHA, Morosoli, 2020), and Prospective Imaging Study of Ageing (PISA; Lupton et al., 2021) completed surveys which included

the Buss-Perry aggression questionnaire (Table 4.3 and 4.4); 1,603 people completed the questionnaire twice. At the time of completion, participants were aged 16 to 73 (Figure 4.2).

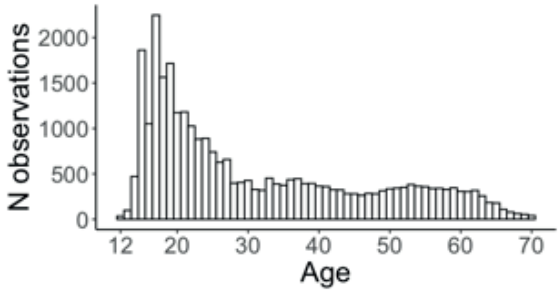


Figure 4.1. The Netherlands: Age distribution

Table 4.3. Australia: Data collection (genotyped individuals)

Study/Year	Observations	Mean age (SD)
16UP (2014-2018)	402	16.34 (0.64)
25UP (2015-2019)	2,052	30.05 (4.31)
GHA (2018-2020)	2,315	40.45 (14.85)
PISA (2016-2020)	2,462	59.98 (6.85)

Table 4.4. Australia: Sample characteristics genotyped individuals

				IRT Aggression	
	***N* subjects**	***N* measures**	**Mean age (*SD*)**	**Mean**	***SD***
Total	5,628	7,231	42.81 (16.71)	-0.01	0.95
Male	1,904	2,385	41.75 (16.84)	0.18	0.90
Female	3,724	4,846	43.33 (16.62)	-0.11	0.96

Note: IRT Aggression= item response theory aggression score, SD= standard deviation.

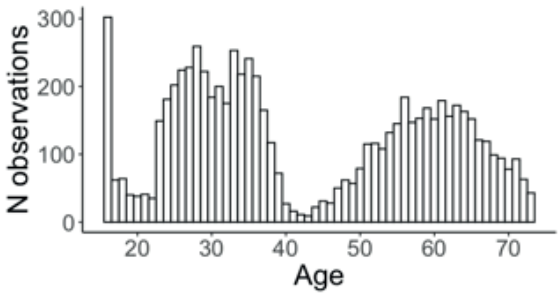


Figure 4.2. Australia: Age distribution

Phenotyping

The Netherlands: All participants completed Achenbach System of Empirically Based Assessment self-report questionnaires (ASEBA; Achenbach et al., 2017), either the Youth Self-Report (YSR; Achenbach & Rescorla, 2001) or the Adult Self-Report (ASR; Achenbach & Rescorla, 2003). In the earlier ANTR surveys, the Young Adult Self Report (YASR) was administered. Surveys in each relevant data collection wave included between 15 and 20 items from the ASEBA aggressive behavior subscale. All items were scored on a three-level scale: 0= never, 1= sometimes, 2= often. Aggression scores were defined separately for each wave of data collection for all NTR participants (i.e. regardless of genotyping status, see Table 4.1) by Item-Response Theory (IRT; Embretson & Reise, 2000) and calculated with the Generalized Partial Credit Model (GPCM) in R (R Core Team, 2017), with the mirt package (Chalmers, 2012). GPCM is an Item Response Theory model, developed to analyze polytomous data. For each wave of data collection, all participants with a maximum of two missing individual items were included in the GPCM. An IRT-aggression score has benefits over a simple sum-score, because it appropriately weights the relative contributions of individual items to a scale with a more favorable distribution, and takes into account missing data. By fitting a separate model for each wave of data collection, aggression scores for each participant are relative to all other participants in that wave of data collection, thereby filtering out potential 'wave' or data collection effects. Because the IRT score for each individual is relative to all other participants in the same wave of data collection, the mean IRT score for each wave is zero. This is reflected in the mean IRT aggression scores at each age (Figure 4.3). The final overall IRT aggression score has a mean of 0.00, and ranges from -1.6 to 4.4, with a standard deviation of 0.86. Only genotyped participants were included for further analysis. The genotyped sample did not differ much from the total sample, with a mean of -0.02, range from -1.6 to 3.5, and a standard deviation of 0.84.

Australia: Aggressive behaviour was measured with the Buss–Perry Aggression Questionnaire. This is a 29-item questionnaire in which participants indicate the extent to which statements are characteristic of them (5-point Likert scale, from "extremely uncharacteristic of me" to "extremely characteristic of me", including some items that needed to be reversed). The questionnaire provides a total sum score and four subscores: Physical aggression, Verbal aggression, Anger, and Hostility. For 107 out of 7,231 observations, missing values on 1 to 6 individual items were

imputed using multivariate imputation via the MICE R-package (van Buuren & Groothuis-Oudshoorn, 2011). IRT aggression scores were calculated with mirt R-package (Chalmers, 2012) within each study (i.e., 16UP, 25UP, GHA, and PISA) for the total aggression score and each of the subscales. Because the IRT score for each individual is relative to all other participants in the same study, the mean score for each study is zero. This is reflected in the mean aggression scores for each age (Figure 4.4). The final overall IRT aggression score has a mean of 0.0, and ranges from -2.9 to 3.8, with a standard deviation of 0.9.

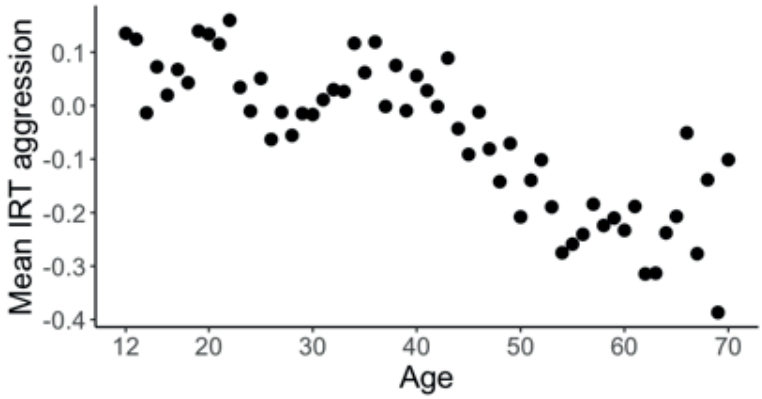


Figure 4.3. The Netherlands: Mean IRT aggression score for each age in the genotyped sample

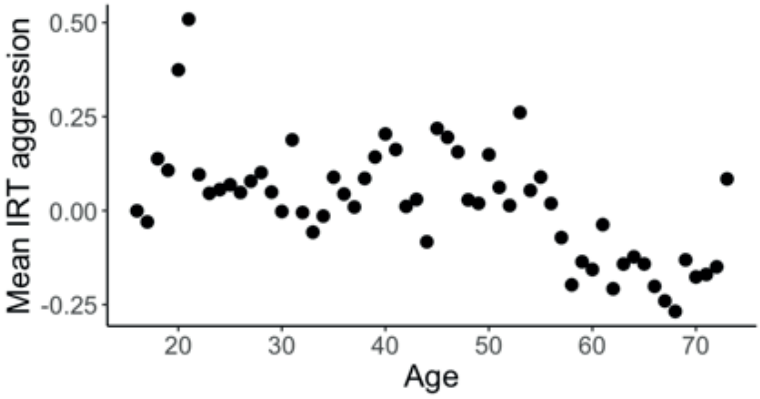


Figure 4.4. Australia: Mean IRT aggression score for each age

Genotype data

The Netherlands: Participants were genotyped on multiple platforms: Affymetrix Axiom, Affymetrix 6.0, Illumina 1M, Illumina 660, Illumina GSA, Perlegen Affymetrix. Samples with call rate <0.90, Plink heterozygosity F< −0.10 or F> 0.10, and inconsistency of X chromosome genotypes with reported gender were excluded. SNPs with MAF <1.0E-6, HWE p-value < 1.0E-6, and/or call rate <0.95 were removed. Genotype data were aligned with the 1000 Genomes reference panel, and filtered for SNPs with allele frequency differences from the CEU population larger than 0.20, palindromic SNPs, and DNA strand issues. DNA Identity By Descent (IBD) state was estimated for all individual pairs using Plink (Purcell et al., 2007) and King (Manichaikul et al., 2010) based on ~10.8k SNPs that all platforms have in common. Samples were removed if IBD did not match expected family relations. CEU population outliers were removed from the data with Smartpca software, based on per platform 1000 Genomes PC projection. Per platform, data were phased using Eagle and imputed to 1000 Genomes with Minimac (Das et al., 2016). The final merged genotype data consist of 12,152,830 SNPs.

Australia: Genotyping was performed on DNA extracted from blood and saliva samples, on Illumina 317K, 370K, 610K, ('1st generation'), GSA, or Core Exome plus Omni-family ('2nd generation') arrays, and GenomeStudio software for genotype calling (Illumina Inc., 200 Lincoln Centre Dr, Foster City, CA 94404). This was followed by imputation from a common SNP set to the 1000 Genomes (Phase 3 Release 5) reference panel, a strategy that allows genotype data from different arrays to be combined. Samples with <97.5% call rate, non-European ancestry (> 6 SD from the mean European-population cluster for PC1 and PC2) or with familial relationships incompatible with those reported by study participants were excluded. Observed markers were cleaned (by batch) for call rate (≥95%); minor allele frequency (≥ 1%); Hardy-Weinberg equilibrium ($p \geq 10^{-6}$), GenCall score (≥ 0.15 per genotype; mean ≥ 0.7) and standard Illumina filters, before integrating batches and re-running relationship and Mendelian checks. Phasing and imputation were carried out at the Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html#!) using the 1000 Genomes Phase 3 Release 5 'mixed population' reference panel, with phasing by SHAPEIT followed by imputation using minimac 3 (Das et al., 2016), '1st generation' and '2nd generation' data were imputed separately due to poor overlap between typed markers. Imputation was based on 277,690 ('1st generation') and 240,297 ('2nd generation') typed markers (passing QC in all relevant batches); and the two were combined after imputation to maximise

sample size, using for each individual the '1st generation' imputation if available, otherwise using the '2nd generation' imputation. This resulted in 9,411,304 SNPs available for analysis, after quality control.

Polygenic score construction

We obtained effect sizes for the association between individual SNPs and aggression from the Ip et al. (2021) GWAMA after omitting the target samples, i.e. analyses were run with no participants from the Netherlands for the Dutch target sample, and no participants from Australia for the Australian target sample (GWAMA sample size for the Netherlands: N_{SNPs}= 7,722,825, $N_{measures}$= 276,268, $N_{individuals}$= 81,259, SNP-h^2 = 3.91 %, SE= 0.42. GWAMA sample size for Australia: N_{SNPs}= 7,762,065, $N_{measures}$= 314,604, $N_{individuals}$= 75,536, SNP-h^2= 3.97%, SE= 0.46). We then computed PGSs for both cohorts with SBayesR V2.03 (Lloyd-Jones et al., 2019), using default settings.

Statistical analyses

To ascertain the viability of predicting adult aggression with the PGS that is based on a discovery in 3- to 18-year-olds, we first model the association between the PGS and aggression in the total sample in the Netherlands and in Australia. Here, IRT aggression is predicted from the PGS, age, age^2, sex, dummy variables for genotyping arrays, and five ancestry-based principal components. We control for the dependence between measures due to relatedness and repeated measures, by adding a random effect for families. Next, we model the effect of the PGS at specific ages. When investigating the effect of PGSs at specific ages, there is a risk that differences in sample size at each age may affect the results. To remedy this, we employ a novel weighted analytic approach in which we make use of more data when looking at specific ages. For each age for which data are available, ages 12 to 70 years in the Dutch context, and ages 16 to 73 years in the Australian context, we model aggression with the package lme4 (Bates, 2015), as a function of the PGS, age, age^2, sex, dummy variables for genotyping arrays, and five ancestry-based principal components. This means a total of 59 analyses in The Netherlands and 58 in Australia. The models are fitted with weights that weight observations at the focus age as 1 and decay for ages further from that age. In this approach, sample size at each age includes ages around the focus age, but to a lesser extent. In other words, data on surrounding ages are included in the analyses, resulting in larger and more comparable sample sizes at different ages. At each age we control for the dependence between measures due to relatedness and repeated measures, by adding a random

effect for families. This captures both dependence due to relatedness and dependence between longitudinal measures. Age covariates are included in the model because a range of ages is still present in each model, albeit with different weights. The mixed effects regression model at each age can be written as:

$$(1) \quad AGG_{mif} = \beta_{0f} + \beta_{\text{Agef}} * \text{Age}_{mi} + \beta_{1\text{-}13} * x_{mi,1\text{-}13} + \epsilon_{if}$$

In this notation, AGG is the aggression outcome of measure *m*, from individual *i* in family *f*, the intercept β_{0f} is a combination of the population level intercept and the family-level deviation of that intercept, β_{Agef} *is the regression estimate for age and the* family-level deviation of that estimate, Age_{mi} is the age of individual *i* at measure *m*, $\beta_{1\text{-}v}$ are the regression estimates for all the fixed effects (the PGS, age, age^{2,} sex, dummy variables for genotyping arrays, and five ancestry-based principal components.), $x_{mi,1\text{-}v}$ are the corresponding observed scores of individual *i* at measure *m for each fixed effect, and* ϵ_{if} is the combined individual and group level error term. In matrix notation, this gives:

$$(2) \quad Y = X \cdot \beta + Z \cdot b + \epsilon$$

In this notation, Y is the matrix of observed responses, X and Z are the design matrices for fixed effects and mixed effects respectively, β is the matrix of unknown fixed parameters, β is the matrix of unknown random parameters, and ϵ is the vector of unobservable model errors. Because we apply weights to observations based on age, the least squares estimates for the fixed effects model parameters in this model are obtained by the following matrix formula, when we subtract away the random effects:

$$(3) \quad \widehat{\beta} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_v \end{bmatrix} = (X' \cdot W \cdot X)^{-1} X' \cdot W \cdot Y$$

Where *Y* is the matrix of observed responses, β_0 to β_v are the regression estimates for the intercept and all fixed effects, *X* is the design matrix, and *W* is the diagonal matrix of weights. The weights we apply to observations in the model are calculated in three steps. In step 1 the vector of ages is weighted as a function of center (age of interest) and shoulder (reflecting kurtosis in the distribution of weights):

(4) $w_1 = 1 - |c - x|^s$

In this notation, w_1 is the weights vector after step 1, x is the vector of ages, c is the center of the weights, and s is the shoulder. In the second step, the weights vector from step 1 is scaled, by applying a min-max scalar:

(5) $w_2=(w_1-min(w_1))/(max(w_1)-min(w_1))$

In this notation, w_2 is the weights vector after step 2, w_1 is the weights vector after step 1. Third, the desired decay is applied to the scaled vector w_2, and the final diagonal matrix of weights is calculated:

(6) $W = diag(w_2^{e^k})$

Where W is the final diagonal matrix of weights, w_2 is the weights vector after step 2 and e^k represents the decay in the distribution of weights. In this approach, we opted for a shoulder of 1.5, and a decay of 25. See Figure 4.5 for an example of the weights for ages 25 and 50. In general, a wider distribution of weights smooths the sample size and age-specific effects more, while a narrower distribution is more sensitive to fluctuations in sample size and age effects. We ran supplemental analyses to test the impact of wider and narrower age weight distributions (see Supplements Figure 4.2 & 4.3).

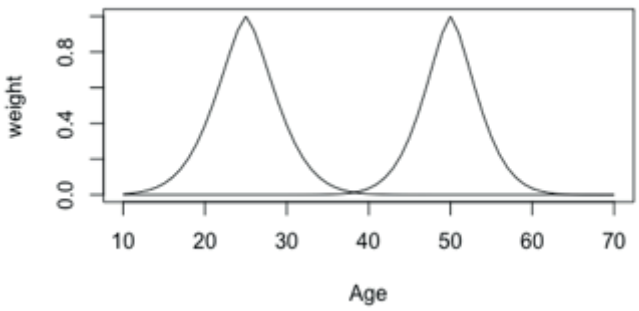


Figure 4.5. Example of weights with centers at ages 25 and 50

All continuous fixed effects and the IRT aggression scores were standardized before the analyses. We employed bootstrapping to assess the robustness of the model estimated standard error. The approach was to sample complete families with replacement from the original data, 100 samples for each age-analysis.

We found that model implied standard errors were slightly underestimated. Therefore, 95% confidence intervals reported in the results were calculated with the bootstrap standard errors. Significance is implied when the empirical (bootstrap) 95% confidence intervals do not intersect zero.

RESULTS

In the total genotyped Dutch sample ($N_{observations}$ = 29,454, $N_{individuals}$ = 13,471), we first analyzed all data together without weighting the data. In this approach, the PGS was significantly associated with aggression, β = 0.05, *SE* = 0.01, *pseudo* R^2 = 0.002, $p < 0.001$. Next, we ran the age-specific age models with rolling weights. These analyses showed that the PGS was significantly related to IRT aggression from age 12 to age 41 (β = 0.04-0.05; Figure 4.6, Supplements Table S4.1). After age 41, the effect of the PGS decreases, with confidence intervals that are close to, or below zero. The highest estimates for the effect of the PGSs are at around age 28. Supplemental analyses with wider and narrower age weights show that the width of the distribution affects the smoothing of the results, meaning that a narrower distribution of weights leads to more defined differences in PGS effects between ages (see Supplements Figure S4.1 & S4.2).

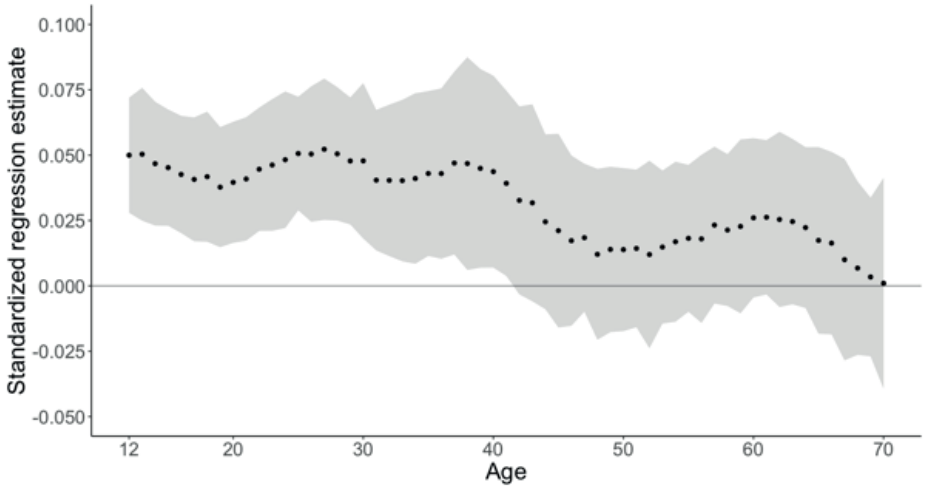


Figure 4.6. Dutch data: standardized regression estimates for the effect of the PGS with bootstrapped 95% confidence intervals (as grey banners)

In the total genotyped Australian sample without rolling weights ($N_{observations}$ = 7,231, $N_{individuals}$ = 5,628), the PGS was significantly associated with adult and adolescent aggression, β = 0.04, *SE* = 0.01, *pseudo* R^2= 0.002, *p* = 0.002. The specific age analyses from the model with rolling weights suggested a different association pattern across age compared to the Dutch sample (See Figure 4.6, Supplements Table S4.2). The results did not indicate a downward trend in the regression estimates. Instead, the PGS was significantly related to aggression at ages 38 to 48 (β = 0.04-0.11).

In the Australian cohort we also investigated whether there are differences in prediction for the four subscales of the Buss Perry Aggression questionnaire (i.e. Physical aggression, Verbal aggression, Anger, and Hostility). The trends are very similar across all subscales. The peak around age 45 is clearly present for all subscales, albeit to a slightly lesser extent in verbal aggression. For figures of the regression estimates from the Buss Perry subscales see Supplements Figure S4.3.

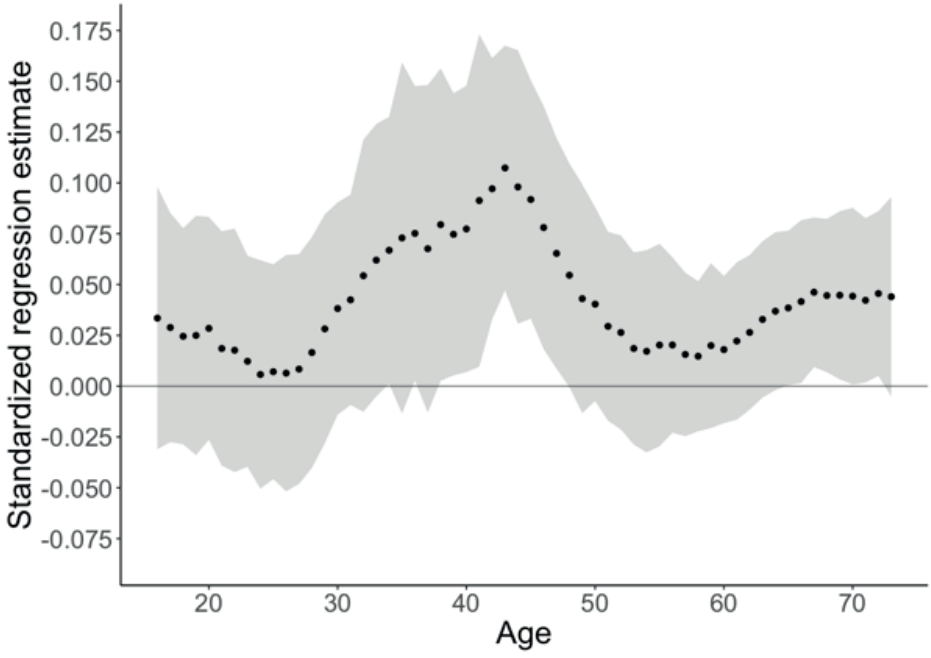


Figure 4.7. Australian data: standardized regression estimates for the effect of the PGS with bootstrapped 95% confidence intervals (as grey banners)

DISCUSSION

In this study we introduce a new method to investigate the effect of a PGS across levels of a continuous moderating variable, in this case age. The approach is to run a linear model for each age in years present in the data. In each analysis, the phenotype information is weighted, with weights that are centered at the focus age and decay further away from that center. The strength of this method is that with each age analysis, information on proximal ages is taken into account, mitigating the risk that sample size differences at different ages will drive the effects we find. We applied this approach to assess the association between childhood aggression PGSs and aggression across the life-course, in two cohorts from The Netherlands and Australia.

In The Netherlands, the PGS was significantly related to aggression at ages 12 to 41 years. The effect of the PGS decreased from age 41 to 70 years. In the smaller Australian cohort, the effect of the PGS was significant at ages 38 to 48. Sample sizes at the peak of the PGS effect in Australia were between *N*= 193 and *N*= 868 (calculated as sum of weights, see Supplements Table S4.2), which was relatively small. Because of these sample sizes, we should interpret this peak in effects with caution; for example, it could be driven by a small number of individuals that are not representative of the population. Effect sizes are small in both cohorts, with just under 0.2% explained variance in the full non-weighted models. We expect that effect sizes will increase as discovery GWAS sample sizes increase.

Although the effects are small, these results are the first indication from a molecular genetics perspective that genetic influences drive part of the continuity and stability of aggressive behavior, and that genetic effects in childhood persist across life, and thus across situations. This suggests that throughout people's lives, most notably in the Dutch context, developmental changes in individuals only slightly impact the polygenic effects on aggression that were apparent in childhood. These results correspond with findings from longitudinal twin studies on the stability of aggression in children (van Beijsterveldt et al., 2003; Eley, Lichtenstein & Moffit, 2003; Porsch et al., 2016), and a twin study on the stability of externalizing psychopathology in adults (Gustavson et al., 2020), where genetic factors accounted for a large part of the stability over time. Van Beijsterveldt et al. (2003) also demonstrated that new genetic influences can contribute

to stability of aggression across different ages. As such, the continuity of polygenic effects across the life-course covers only part of the genetic influences on the stability of aggression across the life-course.

A strength of this study is that we investigate the association between the polygenic score and aggression in two cohorts. This also implies, if we want to compare our results across cohorts, that there are several considerations, including differences in phenotyping, the impact of the discovery data, cultural differences between societies, and genetic differences between populations that may limit the feasibility of a reliable comparison. In the Dutch cohort, phenotyping was done with the ASEBA self-report questionnaires. In the Australian cohort, phenotyping was done with the Buss Perry Aggression self-report questionnaire. Differences in phenotyping are somewhat mitigated by using an IRT latent variable as dependent variable, instead of a sum score. Because different measurement instruments are used, this somewhat limits how we draw conclusions from the differences we found between the cohorts. A large percentage of participants in the discovery childhood aggression GWAMA from which we calculated our PGSs, were also scored with the ASEBA instruments, i.e. self-, parent- and teacher-reports (Ip et al. 2021), which relate directly to the ASEBA self-report questionnaires used in the current Dutch sample. Correlations between self-, parent- and teacher-reports were not very high in Ip and colleagues' GWAMA. Still, we expect slightly greater power in the Dutch cohort compared to the Australian cohort, based on the similarities in the measurement instrument. Another potential source of dissimilarity is that the cohorts were not phenotyped in the same years. In The Netherlands, participants were phenotyped between 1991 and 2014, in Australia between 2014 and 2020, i.e. individuals that were phenotyped at the same age, are often not phenotyped in the same year. Van der Laan et al. (2021-b) show that self-reported aggression in The Netherlands declined from 1991 to 2015. Thus, differences between cohorts may be influenced by time effects. Another difference in phenotyping between the cohorts was that missing data (<20%) were imputed prior to calculating IRT scores, while in the Dutch sample missing data (<20%) were handled by the IRT models when calculating the scores.

The two discovery GWAMAs for the Dutch and Australian cohorts were also not identical. For The Netherlands, discovery data excluded all NTR participants. For Australia, discovery data excluded all Brisbane Longitudinal Twin Study and Prospective Imaging Study of Ageing participants. We opted for this strategy because leaving participants from both cohorts out of the

discovery data would mean an unnecessary decrease in sample size. The differences in the discovery GWAMAs means that the PGSs in both samples are calculated based on overlapping, but not identical information. This may have resulted in slight differences in association between cohorts, although respectively only 5 and 17 percent of the Dutch and Australian discovery samples was not shared.

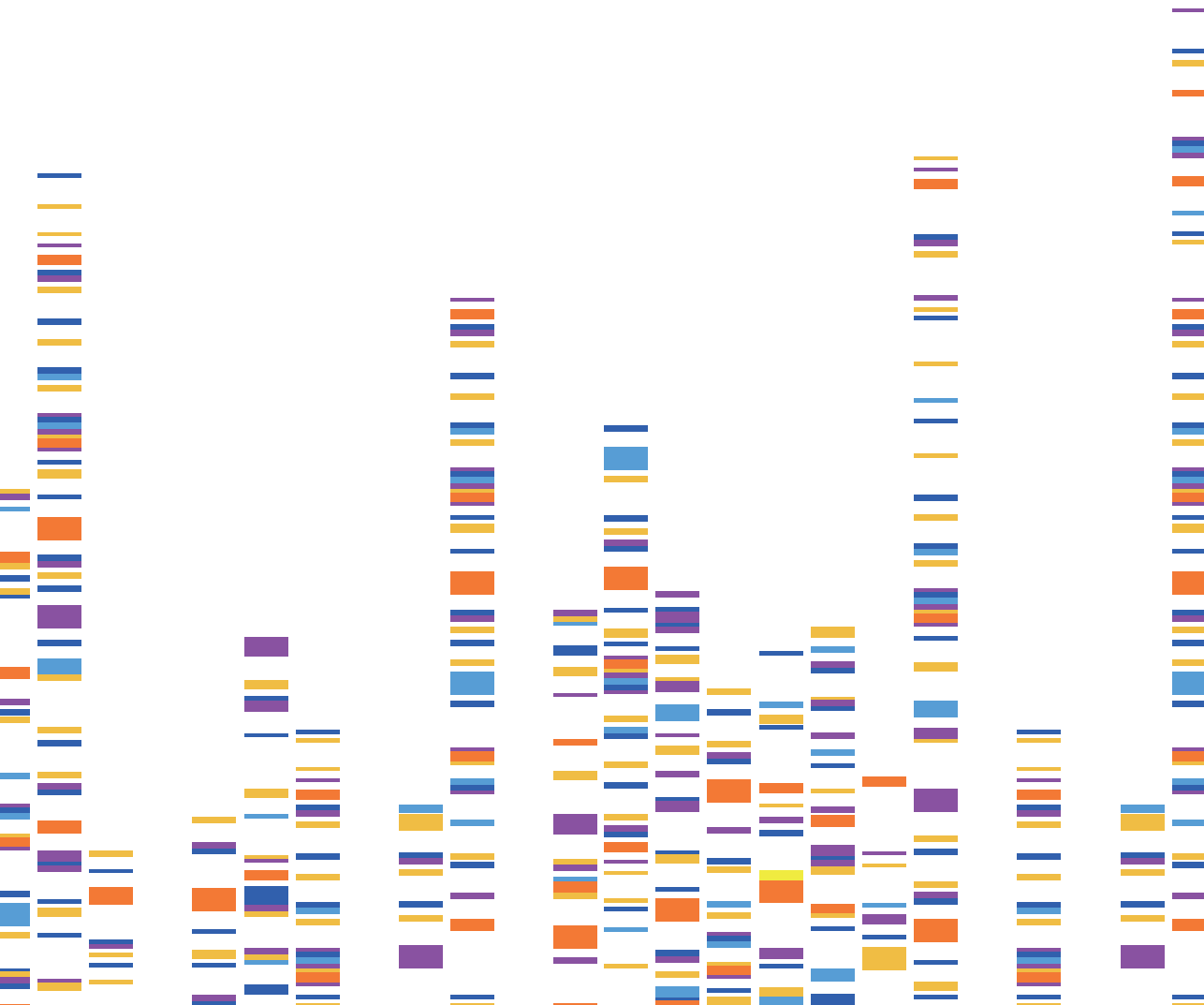
More generally, even if phenotyping is similar, cultural and genetic differences between populations can affect the magnitude, and thereby the merit, of PGS predictions. Cultural norms may influence aggression directly, and thereby moderate the effect of PGSs on aggression. One way to measure the effects of cultural differences on traits is by assessing generalizability of the measurement instruments via Confirmatory Factor Analysis (CFA), using the framework of measurement invariance (Millsap, 2011). By assessing measurement invariance, we can test whether we measure the same underlying psychopathological trait when studying different societies. Because the Australian and Dutch cohorts did not phenotype by the same instruments, we cannot test measurement invariance directly in our sample. However, measurement invariance for the ASEBA self-report questionnaires is well documented. CFA of the eight-syndrome structure of the youth self-report, originally derived from a U.S. general population sample, plus clinically referred youths from Australia, England, and the United States (Achenbach & Rescorla, 2001), fits YSR data from a wide range of societies (Ivanova et al., 2007). Fit indices were almost identical between Australia, RMSEA= 0.042 and The Netherlands, RMSEA= 0.040. Ivanova and colleagues (2015) also investigated generalizability of the eight-syndrome structure of the adult self-report (ASR) in 29 societies. Although The Netherlands and Australia were not included, model fit was good for all samples, with fit indices very similar to those found in the YSR CFAs. Generalizability of the Buss Perry aggression questionnaire has been less extensively studied, but validation is well documented for young Western adults (see for an overview Gerevich et al., 2007). Generalizability has been questioned for older and more diverse samples, but the four-factor structure (Physical aggression, Verbal Aggression, Hostility, and Anger) did replicate in a sample of Chilean students (Valdivia-Peralta et al., 2014), a slimmed down, translated 12 item version (Bryant & Smith, 2001) replicated well in a sample of Hong Kong Chinese (Maxwell, 2007), and all factors except Anger replicated in Hungarian adults with a mean age of 46.6 years (Gerevich et al., 2007).

Another source of differences between the population cohorts may be genetic heterogeneity. Based on genetic marker data, there are several ways to assess the comparability of samples from different populations. In an early comparison based on DNA tandem repeat polymorphisms from The Netherlands and Australia, Sullivan et al. (2006) estimated the fixation index (F_{ST}) – the percentage genetic variability attributable to genetic differences between cohorts – to be only 0.3%. The empirical variability between the Australian and Dutch cohorts was in fact smaller than for a combination of European samples (FST= 0.8%; Rosenberg, 2021). More recently, Beck and colleagues (2019) studied interpopulation stratification between cohorts from The Netherlands and Australia, by analyzing genome wide SNP data. Negligible interpopulation stratification was confirmed by visualizing uncorrelated principal components and FST estimations of 0.05%. These findings should not be surprising, given the colonization history of Australia via immigrants from the UK, who are genetically very similar to Dutch individuals, as previously described with similarities in Y-chromosome haplogroups (Rosser et al., 2000).

At the moment, this genetic similarity between cohorts included in genomic studies is the rule rather than the exception. Around 78% of individuals included in GWAS are from European ancestry (EA; Buniello et al., 2019). The lack of individuals from non-EA populations included in GWAS, means that it is often unclear to what extent genetic effects generalize to diverse populations. Carlson and colleagues (2013) demonstrated that in non-EA populations, most GWAS-identified variants have allelic associations in the same direction as in EA populations, with none showing a statistically significant effect in the opposite direction. However, 25% of tagSNPs had significantly different effect sizes in at least one non-EA population, most frequent in African Americans, with all differential effects diluted towards zero. Thus, associations between the PGS based on the Ip et al. (2021) discovery and aggression might be weaker in non-EA populations, as has been seen for other traits, such as obesity/BMI (Domingue et al., 2014; Belsky et al., 2013; Ware et al., 2017), height (Ware et al., 2017), educational attainment (Domingue et al., 2015; Ware et al., 2017; Lee et al., 2018), schizophrenia (Vilhjálmsson et al., 2015; Ware et al., 2017; Vassos et al., 2017), and breast cancer (Ho et al., 2020). This means that advances through GWAS and PGS studies tended to benefit EA populations more than non-EA populations, especially when predicting health outcomes (Martin et al., 2019). The overrepresentation of EA participants in GWAS is partly because admixed populations were long considered inconvenient in gene discovery

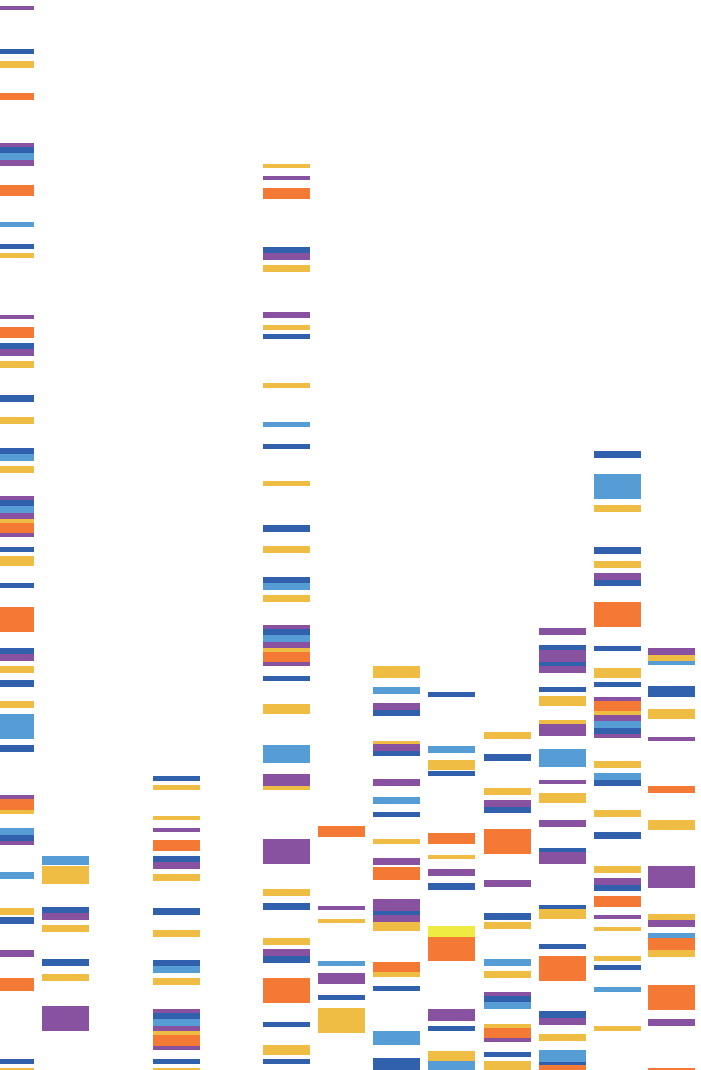
studies, as this led to population stratification issues. However, due to advances in GWAS methods, populations with mixed genetic backgrounds can now be included in GWAS to obtain accurate estimates of SNP effects, boost power, and improve fine-mapping of effects by leveraging linkage disequilibrium differences (Asimit et al., 2016; Atkinson et al., 2020). Beside the general benefits to gene discovery studies, the inclusion of diverse genetic backgrounds will improve our understanding of genetic liability across diverse populations, as demonstrated for example in a study of glycemic traits (Chen et al., 2021), where 30% of the participants were of non-European ancestry.

In summary, we investigated the continuity of polygenic effects on aggression across the life-course, in two cohorts from The Netherlands and Australia, with a novel weighted mixed effects regression approach. Our results suggest that the same genetic factors that explain part of the individual differences in aggression in childhood, also explain individual differences in adolescents and adults. The new method we employed shows promise in modeling genetic effects across levels of a continuous moderating variable, in way that smooths any possible effects due to sample size differences between those levels. The possibilities of reliably comparing results between the Dutch and Australian cohorts were limited because of differences in phenotyping and GWAMA discovery samples. When studying genetic liability in different populations, there are two main considerations: cultural/environmental differences and genetic differences. If we are interested in studying differences in genetic effects between populations with different cultural norms and environments, the optimal design is to look at populations with similar genetic backgrounds. For this to work, phenotyping has to be standardized across populations. To better understand genetic effects in diverse genetic populations, we need gene discovery studies that include diverse populations, so that predictions in non-EA populations are not dependent on EA discovery samples.



5

Een Tweelingstudie naar Indicatoren van Genetische en Culturele Transmissie



ABSTRACT

In deze tweelingstudie is de rol van genetische en culturele transmissie bij intergenerationele continuïteit van norm-overschrijdend gedrag (NoG) onderzocht. Op grond van de overeenkomsten binnen 3.982 Nederlandse tweelingparen van 13 tot 17 jaar werd de relatieve invloed van genetische factoren (G), gedeelde omgeving (C) en unieke omgeving (E) op NoG geschat. Culturele transmissie, het doorgeven van kennis, normen en waarden, vormt deels de gedeelde omgeving van kinderen die opgroeien in hetzelfde gezin. We vonden geen significante invloed van gedeelde omgeving, en daarmee geen evidentie voor culturele transmissie. Genetische invloeden verklaarden 60 procent van de individuele verschillen in norm-overschrijdend gedrag van 13- tot 17-jarigen. Deze resultaten suggereren dat intergenerationele continuïteit bij norm-overschrijdend gedrag op deze leeftijd vooral gedreven wordt door genetische transmissie.

Gebaseerd op:
van der Laan, C.M., van de Weijer, S.G.A., Nivard, M.G., & Boomsma, D.I. (2019). Een tweelingstudie naar indicatoren van genetische en culturele transmissie. *Tijdschrift Voor Criminologie*, *61*(2), 148–164.

English summary

In the present study, the role of genetic and cultural transmission in intergenerational continuity of rule-breaking behavior (RBB) was investigated. Based on the resemblance within 3,982 Dutch twin pairs, aged 13 to 17 years, the relative importance of genetic (G), shared environmental (C), and unique environmental (E) influences on RBB was estimated. Cultural transmission, the process of passing on knowledge, norms and values, can lead to similarities within families, and forms part of the shared environment of children growing up in the same family. The authors found no evidence for shared environmental influences, and consequently no indication of a role for cultural transmission. Genetic influences explained 60 percent of the variance in rule-breaking behavior at age 13 to 17, implying that intergenerational continuity at this age is mainly driven by genetic transmission.

Based on:
van der Laan, C.M., van de Weijer, S.G.A., Nivard, M.G., & Boomsma, D.I. (2019). Een tweelingstudie naar indicatoren van genetische en culturele transmissie. *Tijdschrift Voor Criminologie*, *61*(2), 148–164.

EEN TWEELINGSTUDIE NAAR INDICATOREN VAN GENETISCHE EN CULTURELE TRANSMISSIE

Kinderen van criminele ouders hebben een verhoogd risico om zelf ook crimineel en antisociaal gedrag te vertonen. Besemer en collega's (2017) stelden, op basis van een meta-analyse van 25 internationale publicaties, dat het risico dat kinderen van criminele ouders zelf crimineel gedrag lieten zien 2,4 maal hoger was dan bij kinderen zonder criminele ouders. Wanneer gecontroleerd werd voor verschillende risico- en beschermende factoren, was het risico op crimineel gedrag nog steeds 1,8 maal hoger dan bij kinderen zonder criminele ouders. Deze patronen werden ook in Nederland gevonden (o.a. Junger e.a., 2013; Bijleveld & Wijkman, 2009). Deze en eerdere studies wijzen in de richting van intergenerationele continuïteit van antisociaal en crimineel gedrag. Er is minder bekend over de onderliggende mechanismen die deze patronen zouden kunnen verklaren. In de criminologie worden verschillende theorieën aangehaald als mogelijke verklaringen. Hierbij wordt vooral gekeken naar omgevingsfactoren, terwijl er relatief weinig aandacht uitgaat naar genetische transmissie.

De overerving van genetische factoren, oftewel genetische transmissie, verloopt via de wetten van Mendel. Beide ouders geven de helft van hun genetische materiaal door aan hun nageslacht. Dat heeft verschillende belangrijke implicaties. Kinderen met dezelfde ouders lijken meer op elkaar dan kinderen die niet dezelfde ouders hebben. Tegelijkertijd kunnen binnen hetzelfde gezin kinderen voorkomen die fors van elkaar verschillen, omdat ze andere varianten erfden van hun ouders (Boomsma e.a., 2002). Omgevingsfactoren kunnen ook leiden tot verschillen en overeenkomsten binnen gezinnen. Ze kunnen worden gedeeld door kinderen die opgroeien in hetzelfde gezin, en ertoe leiden dat kinderen op elkaar lijken los van hun genetische verwantschap. Daarnaast kunnen omgevingsfactoren uniek zijn voor ieder kind en bijdragen aan de verschillen tussen broers en zussen. De gedeelde omgevingsfactoren kunnen onder andere ontstaan als er sprake is van 'culturele transmissie' (Keller e.a., 2009). Bij verticale culturele transmissie worden kennis, normen en waarden doorgegeven van de ene generatie op de volgende (Cavalli-Sforza & Feldman, 1981). Culturele transmissie van ouders naar kinderen kan leiden tot gelijkenissen tussen de kinderen in hun gedrag die niet worden verklaard door genetische transmissie. Belangrijk hier is dat wanneer geen rekening gehouden wordt met de mogelijkheid van genetische transmissie, de evidentie voor culturele

transmissie onzeker is. Wanneer culturele transmissie een rol speelt bij intergenerationele continuïteit van gedrag, zullen kinderen van dezelfde ouders meer op elkaar lijken dan kinderen die niet dezelfde ouders hebben.

In deze studie maken we gebruik van een klassiek tweelingendesign. We gebruiken dit design om te schatten welk deel van de individuele verschillen in norm-overschrijdend gedrag (NoG) – waarbij regels en normen thuis of op school worden overtreden – kan worden verklaard door genetische factoren (G), gedeelde omgevingsfactoren (C, van het Engelstalige *common environment*, dat de invloed weergeeft van omgevingsfactoren die gedeeld (*shared*) zijn) en unieke omgevingsfactoren (E, van het Engelstalige *environment*). De bijdrage van genetische factoren impliceert genetische transmissie, de bijdrage van gedeelde omgevingsfactoren kan culturele transmissie impliceren.

Eerder onderzoek

In onderzoek naar crimineel en antisociaal gedrag is meerdere keren gevonden dat erfelijke aanleg een rol speelt. Al in 1984 verscheen een adoptiestudie van Mednick, Gabrielli en Hutchings naar crimineel gedrag van 14.427 Denen die op jonge leeftijd geadopteerd werden, en hun adoptie- en biologische ouders. De kinderen werden meteen uit het ziekenhuis geadopteerd (25 procent) of overgebracht naar een weeshuis, waardoor alleen de prenatale omgeving werd gedeeld met de biologische moeder. De onderzoekers vonden geen significante relatie tussen crimineel gedrag van geadopteerde personen en hun adoptieouders. Wel was er een correlatie in de mate van betrokkenheid bij vermogensdelicten tussen de geadopteerde personen en hun biologische ouders. Dit suggereert dat genen een rol spelen bij de intergenerationele continuïteit van crimineel gedrag, en dat culturele transmissie geen of een kleine rol speelt. Voor geweldsdelicten werd geen significante correlatie tussen de geadopteerde personen en hun biologische ouders gevonden. Hier geven de auteurs geen verdere verklaring voor. Uit, onder andere, tweelingstudies weten we echter dat ook bij agressief gedrag genen een belangrijke rol spelen (Veroude e.a., 2016).

Er is een aantal tweelingstudies verschenen waarbij gekeken werd naar antisociaal of crimineel gedrag. Ferguson (2010) voerde een meta-analyse uit op resultaten uit tweelingonderzoek naar antisociaal gedrag, variërend van externaliserend probleemgedrag tot gewelddadig gedrag. In de meta-analyse zijn 38 publicaties opgenomen, wat leidde tot een totale steekproef van 96.918 personen. Uit zijn onderzoek volgde dat verschillen in antisociaal

gedrag gemiddeld voor 56 procent te verklaren zijn door genetische factoren, voor 11 procent door gedeelde omgevingsfactoren en voor 31 procent door unieke omgevingsfactoren. Ferguson (2010) keek ook naar modererende effecten van leeftijd (kinderen tot 12 jaar, adolescenten van 12 tot 18 jaar en volwassenen vanaf 18 jaar), meetmethode en sekse. De relatieve invloed van genetische factoren en unieke omgeving hing samen met leeftijd, waarbij een hogere leeftijd leidde tot een kleinere relatieve invloed van genetische factoren en een grotere invloed van unieke omgevingsfactoren. Ook de gebruikte meetmethode hing samen met de effectgroottes, waarbij engere definities, zoals de antisociale persoonlijkheidsstoornis volgens de DSM IV, leidden tot kleinere genetische effectgroottes. Breder gedefinieerde concepten, zoals agressief gedrag, leidden tot grotere genetische effectgroottes. Ferguson (2010) vond geen significante invloed van sekse op de schattingen van genetische en omgevingsfactoren. Veroude en collega's (2016) publiceerden een review van genetische studies naar agressief gedrag. De uitkomstmaat van de opgenomen studies bleef echter niet beperkt tot agressief gedrag, ook norm-overschrijdend en crimineel gedrag werden meegenomen. De review bevat 40 tweelingstudies, met steekproeven variërend van 100 tot 36.877 personen. De conclusies waren vergelijkbaar met die van Ferguson (2010). Schattingen van genetische invloeden varieerden rond een gemiddelde van 50 procent, en schattingen van gedeelde omgevingsinvloeden varieerden van 0 tot 35 procent. Veroude en collega's (2016) vonden, in tegenstelling tot Ferguson (2010), weinig aanwijzingen voor een effect van leeftijd op de schattingen van de invloed van genen en omgeving. In de betreffende studies werd vooral gekeken naar adolescenten, waardoor weinig variatie in leeftijden aanwezig was. Daarnaast vonden Veroude en collega's (2016), net als Ferguson (2010), weinig aanwijzingen voor sekseverschillen in de relatieve invloed van genen en omgeving.

In Nederland deden Bartels, Derks en Boomsma (2005) onderzoek naar genetische en omgevingsinvloeden op individuele verschillen in NoG van kinderen van 7, 10 en 12 jaar oud. De ouders rapporteerden over het gedrag van hun kinderen. Bij 7-, 10- en 12-jarige meisjes werd ongeveer 40 procent van de variantie verklaard door genetische invloeden, en rond de 35 procent door gedeelde omgevingsinvloeden. Bij jongens leken genetische invloeden een iets grotere rol te spelen, met geschatte erfelijkheid van 50 tot 60 procent; gedeelde omgevingsinvloeden verklaarden bij jongens ongeveer 30 procent van de variantie.

De huidige studie

Het doel van de huidige studie is om een beeld te krijgen van het belang van genetische en culturele transmissie bij intergenerationele continuïteit van NoG bij jongeren. Hierbij zijn de schattingen van erfelijkheid en gedeelde omgeving een indicator voor genetische en culturele transmissie. We kijken in dit onderzoek naar zelfgerapporteerd NoG van 13- tot 17-jarigen. Onderzoek naar crimineel gedrag gedurende de levensloop laat zien dat dit gedrag zich vooral manifesteert tijdens de adolescentie (Van der Laan & Goudriaan, 2016; Matthews & Minton, 2017). Deze piek in crimineel en antisociaal gedrag maakt deze leeftijdscategorie extra interessant voor onderzoek naar genetische en culturele transmissie van NoG.

METHODE

Deelnemers

Tweelingen in dit onderzoek staan ingeschreven bij het Nederlands Tweelingen Register (NTR: www.tweelingenregister.org/). De NoG-data zijn verzameld via zelfrapportage. Toen de tweelingen ongeveer 14 jaar oud waren, werd hun ouders toestemming gevraagd om de kinderen een vragenlijst toe te sturen. Vervolgens ontvingen de tweelingen ieder een eigen uitnodiging om mee te doen aan het onderzoek en een eigen vragenlijst. De steekproef bestaat uit 5.218 tweelingparen (3.982 complete paren en 1.236 incomplete paren) van 13 tot 17 jaar. Van de tweelingparen zijn 1.884 paren eeneiig, of monozygoot (MZ), en 3.334 twee-eiig, of dizygoot (DZ). Zygositeit bij tweelingen van gelijk geslacht werd bepaald door een synthese van alle beschikbare antwoorden op de vragenlijsten, ingevuld door de tweelingen zelf en/of andere beoordelaars, of door het meten van bloedgroepen of DNA-markers. Van alle tweelingen van gelijk geslacht in dit onderzoek is bij 28 procent de zygositeit op grond van DNA-onderzoek vastgesteld. Bij de overige tweelingen geldt dat de zygositeit is bepaald op basis van de informatie uit de longitudinale vragenlijsten. Geschat wordt dat in ongeveer 95 procent van de gevallen zygositeit correct wordt vastgesteld (Rietveld e.a., 2000). Tweelingen van ongelijk geslacht zijn uiteraard altijd twee-eiig. De tweelingparen zijn verdeeld in vijf groepen: MZ mannen, DZ mannen, MZ vrouwen, DZ vrouwen en DZ ongelijke sekse. Tabel 5.1 geeft de informatie over de groepsgroottes.

Tabel 5.1. Steekproefgrootte, gemiddelde score NoG en leeftijd

Groep	Tweeling	N	M NoG_IRT	SD NoG_IRT	M NoG_SOM	SD NoG_SOM	M Leeftijd	SD Leeftijd
Totaal		9.200	-0,01	0,80	3,01	2,56	14,66	0,64
Jongens		4.063	0,12	0,79	3,37	2,59	14,67	0,63
Meisjes		5.137	-0,09	0,79	2,80	2,51	14,65	0,65
MZM	1	679	0,10	0,80	3,28	2,51	14,68	0,65
	2	666	0,09	0,82	3,28	2,73	14,68	0,64
DZM	1	611	0,16	0,77	3,48	2,47	14,65	0,60
	2	606	0,18	0,80	3,57	2,70	14,66	0,62
MZV	1	1.050	-0,16	0,77	2,59	2,43	14,68	0,66
	2	1.038	-0,11	0,78	2,72	2,41	14,68	0,64
DZV	1	801	-0,06	0,78	2,88	2,53	14,60	0,64
	2	768	-0,03	0,79	2,98	2,59	14,60	0,65
DZOmv	1 (m)	1.501	0,11	0,78	3,34	2,53	14,69	0,62
	2 (v)	1.480	-0,07	0,80	2,88	2,55	14,67	0,64

Noot: MZM= monozygote jongens, DZM= dizygote jongens, MZV= monozygote meisjes, DZV= dizygote meisjes, DZOmv= dizygote ongelijk geslacht, m= jongen, v= meisje, NoG_IRT= IRT-score norm-overschrijdend gedrag, NoG_SOM= somscore norm-overschrijdend gedrag.

Meetinstrument

In deze studie werd een gestandaardiseerde zelfrapportagevragenlijst voor jongeren gebruikt (YSR; Verhulst e.a., 2013). Dit is een vragenlijst voor jongeren van 12 tot 18 jaar, met 118 stellingen over emotionele en gedragsproblemen. De stellingen hebben drie antwoordmogelijkheden: 0=helemaal niet; 1=een beetje of soms; 2=duidelijk of vaak. De norm-overschrijdend gedrag schaal bestaat uit vijftien stellingen (zie Tabel 5.2). Om tot een totaalscore voor NoG te komen is gebruik gemaakt van een Item Response Theory (IRT)-model (zie o.a. Embretson & Reise, 2000). Bij IRT wordt niet aangenomen dat elk item uit de schaal even informatief is voor het meten van NoG, maar wordt berekend hoe groot de kans is dat – in dit geval – iemand 0, 1 of 2 antwoordt, in relatie tot de onderliggende ware score voor NoG. Hieruit komt een score die te vergelijken is met een factorscore waarbij de scores variëren rond een gemiddelde van 0. Dit leidt tot een score voor NoG die beter verdeeld is dan een simpele somscore (zie Figuur 5.1 en 5.2). Uit de IRT-gewichten voor de verschillende items kan worden afgeleid welke items meer of minder van belang zijn bij het meten van NoG. Zo lijkt, op basis van het IRT-gewicht, een gebrek aan schuldgevoel slechts marginaal bij te dragen aan de NoG-score. Over het algemeen is er een patroon te zien

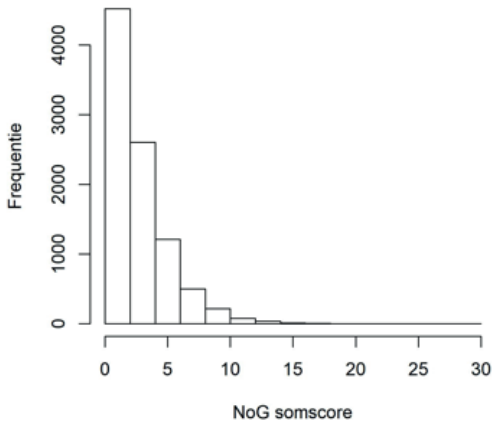
waarbij de 'serieuzere' items belangrijker zijn bij het meten van NoG, met druggebruik als zwaarstwegend. Voor deelnemers werd een maximum van twee missende antwoorden op de vragen over NoG toegestaan.

Tabel 5.2. Items YSR norm-overschrijdend gedrag

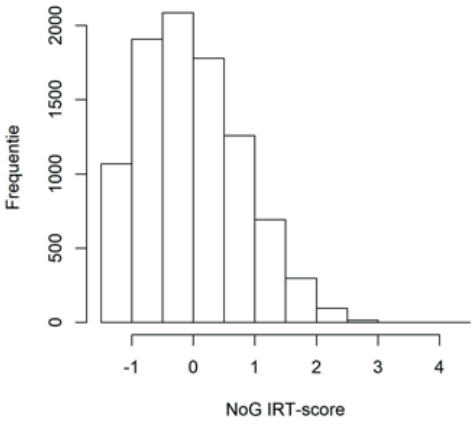
Items	IRT-gewicht	N missend
Alcohol consumeren zonder toestemming ouders	0,618	25
Gebrek aan schuldgevoel	0,154	36
Overtreden van regels	0,651	47
Problematische vrienden	0,496	50
Liegen of bedriegen	0,602	49
Voorkeur voor oudere vrienden	0,410	23
Weglopen van huis	0,666	9
Brand stichten	0,638	13
Stelen uit huis	0,679	21
Stelen buitenshuis	0,648	18
Vloeken	0,566	10
Overmatige gedachten aan seks	0,478	26
Roken	0,692	17
Spijbelen	0,733	52
Druggebruik	0,802	63

Analyses

Tweelingstudies maken gebruik van de verschillen in genetische verwantschap tussen een- (MZ) en twee-eiige (DZ) tweelingen om het relatieve belang van genetische factoren (G), gedeelde omgevingsfactoren (C) en unieke omgevingsfactoren (E) te onderscheiden. MZ tweelingen ontstaan wanneer een bevruchte eicel in tweeën splitst, en delen daarom 100 procent van hun erfelijk materiaal. DZ tweelingen ontstaan na een dubbele ovulatie bij de moeder, en delen gemiddeld 50 procent van hun genetisch materiaal, zoals ook het geval is bij andere broers en zussen. De mate waarin MZ tweelingen van elkaar verschillen, is toe te schrijven aan unieke omgevingsfactoren, terwijl bij DZ tweelingen ook genetische verschillen een rol spelen. Overeenkomsten, vaak uitgedrukt in correlaties, tussen MZ en DZ tweelingen zijn een functie van genetische factoren en gedeelde omgevingsfactoren. Tabel 5.3 geeft een overzicht van de mogelijke conclusies die op grond van de vergelijking tussen correlaties van MZ en DZ tweelingparen getrokken kunnen worden.



Figuur 5.1. Verdeling somscore NoG



Figuur 5.2. Verdeling IRT-score NoG

Tabel 5.3. Tweelingcorrelaties en gerelateerde conclusies

Tweelingcorrelaties	Conclusie
r(MZ)=r(DZ)=0	Alleen unieke omgevingsinvloeden, want kinderen uit hetzelfde gezin lijken niet op elkaar
r(MZ)=r(DZ)>0 en r(MZ)=r(DZ)<1	Unieke omgevingsinvloeden & gedeelde omgevingsinvloeden, want MZ tweelingen lijken niet meer op elkaar dan DZ tweelingen, en de correlaties zijn kleiner dan 1
r(MZ)=2*r(DZ) en r(MZ)<1	Unieke omgevingsinvloeden & genetische invloeden, want MZ tweelingen lijken meer op elkaar dan DZ tweelingen, en de 2x zo grote gelijkenis suggereert dat gedeelde omgeving geen rol speelt
r(MZ)<2*r(DZ) en r(MZ)<1 en r(DZ)<1	Unieke omgevingsinvloeden & gedeelde omgevingsinvloeden & additief genetische invloeden, want MZ tweelingen lijken meer op elkaar dan DZ tweelingen, maar niet 2x zoveel

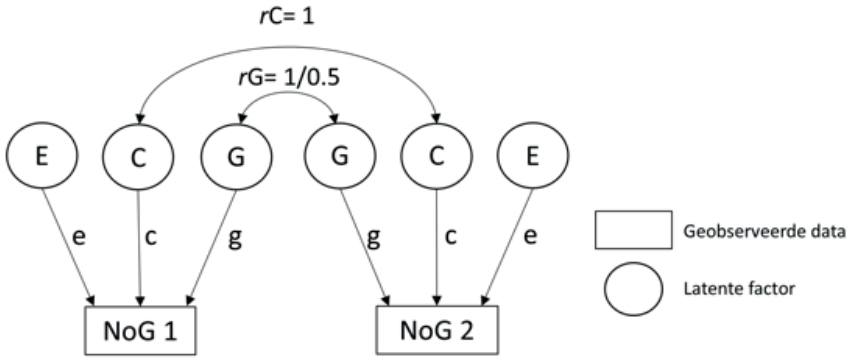
Noot: r(MZ): correlatie tussen monozygote tweelingparen; r(DZ): correlatie tussen dizygote tweelingparen.

Als we voor een bepaald gedrag (in de genetica vaak aangeduid met fenotype) zoals NoG schrijven: NoG=gG+cC+eE, staat NoG voor de geobserveerde uitkomst, en zijn G, C en E de niet-geobserveerde of latente variabelen. Deze notatie wordt ook vaak gebruikt in *structural equation modeling* (SEM). Wanneer de latente factoren gestandaardiseerd zijn (en dus een variantie van 1 hebben), en hun invloed wordt weergegeven door parameters g, c en e, kan de variantie in de uitkomstmaat, in dit geval NoG, worden geschreven als:

$$Var(NoG)=g^2+c^2+e^{2,}$$

waarbij g^2, c^2 en e^2 dus sommeren tot de totale variantie in NoG. De relatieve invloed van de individuele factoren op NoG is dan een proportie van de totale variantie. Voor de relatieve invloed van de genetische factoren geldt daarom: $h^2=g^2/Var(NoG)$, waarbij h^2 staat voor erfelijkheid (h^2 is afgeleid van het Engelstalige *heritability*). De geschatte erfelijkheid kan, naast de pure genetische invloeden ('main effects'), ook variantie bevatten die het gevolg is van niet-gemodelleerde, niet-lineaire genetische invloeden, en van gen-gedeelde omgevingsinteractie (GxC). Voor de relatieve invloed van de gedeelde omgeving geldt: $c^2/Var(NoG)$. De gedeelde omgeving kan, naast omgevingsinvloeden, ook variantie bevatten die het gevolg is van 'assortative mating', waarbij ouders, en daarmee DZ tweelingen, genetisch meer op elkaar lijken dan bij 'random mating' verwacht zou worden. Ook kan de gedeelde omgeving gen-gedeelde omgevingscorrelatie bevatten (rGC). Voor de relatieve invloed van de unieke omgeving geldt: $e^2/Var(NoG)$. De unieke omgeving omvat alle omgevingsinvloeden die leiden tot verschillen binnen tweelingparen, en eventuele gen-unieke omgevingsinteractie (GxE). Met SEM kunnen, op grond van de covariantiestructuren, *maximum likelihood* (ML)-schattingen verkregen worden van de erfelijkheid, de relatieve invloed van de gedeelde omgeving en de relatieve invloed van de unieke omgeving. In Figuur 5.3 staat een voorbeeld van een klassiek tweelingmodel weergegeven. Als we dit model vertalen naar het correlatiepatroon in MZ en DZ tweelingparen, dan zijn de verwachtingen voor de correlaties: $r(MZ)=h^2+c^2$ en $r(DZ)=½\ h^2+c^2$. Een eerste schatting van erfelijkheid kan worden verkregen door het verschil in correlaties tussen MZ en DZ tweelingen te verdubbelen: $h^2=2*(r(MZ)-r(DZ))$. De bijdrage van gedeelde omgeving wordt geschat door $2*r(DZ)-r(MZ)$. Tot slot kan de bijdrage van unieke omgevingsinvloeden worden geschat door: $1-r(MZ)$. De genetische effecten zijn altijd 1 gecorreleerd in MZ paren, bij de DZ paren wordt uitgegaan van een additief genetisch model, waarbij kleine effecten

van verschillende genen sommeren tot de totale genetische invloeden en de verwachting voor de genetische correlatie gemiddeld 0,5 is. Bij DZ paren van ongelijk geslacht kan worden geschat of de genetische correlatie kleiner is dan 0,5, wat zou betekenen dat bij jongens en meisjes verschillende genen tot expressie komen.



Figuur 5.3. Klassiek tweelingmodel
Noot: Var(NoG)=$g^2+c^2+e^2$
Uit de paden volgen ook de verwachte covarianties tussen tweelingen: cov(MZ)=g^2+c^2, cov(DZ)=$½g^2+c^2$. G, C en E zijn latente factorscores (genetische factorscore, gedeelde omgeving, unieke omgeving).

Gemiddeldes, standaarddeviaties en correlaties tussen tweelingen zijn berekend in *R* (version 3.4.2; R Core Team, 2017). Voor het schatten van de paden van de latente factoren, g, c en e, zoals weergegeven in Afbeelding 5.3, werd de SEM-software OpenMx (Neale e.a., 2016) in *R* gebruikt. SEM staat toe om verschillende modellen te vergelijken, bijvoorbeeld modellen met en zonder genetische factoren. Op basis van een likelihood-ratio test kan het meest waarschijnlijke model, gegeven de geobserveerde data, worden gekozen.

Door de *fit* van verschillende geneste modellen te vergelijken wordt antwoord gezocht op verschillende vragen: Is er een significante invloed van genetische factoren en van gedeelde omgevingsfactoren? Zijn er sekseverschillen in de relatieve invloed van genetische factoren, gedeelde omgeving en unieke omgeving (kwantitatieve sekseverschillen)? Zijn de genetische factoren die bij jongens en meisjes tot expressie komen gelijk (kwalitatieve sekseverschillen)? De parameters die geschat worden in de modellen zijn de gemiddelde NoG-scores, het lineaire effect van leeftijd op NoG, de genetische invloeden, de gedeelde omgevingsinvloeden, de unieke

omgevingsinvloeden en de genetische correlatie tussen DZ tweelingen van ongelijk geslacht. Alle parameters worden eerst vrij geschat voor jongens en meisjes. Vervolgens wordt gekeken naar modellen waarbij parameters gelijk werden gesteld voor jongens en meisjes, naar modellen waarbij de invloed van genetische factoren en/of gedeelde omgeving gelijk zijn aan nul, en naar een model waarbij de genetische correlatie tussen tweelingen van ongelijk geslacht gelijk is aan 0,5. Dit is een test voor kwalitatieve sekseverschillen. Kwalitatieve sekseverschillen duiden erop dat verschillende genen tot expressie komen bij jongens en meisjes. Kwalitatieve sekseverschillen hoeven niet te leiden tot een verschil in erfelijkheid tussen jongens en meisjes. Ook al spelen verschillende genen een rol, de mate van variantie die deze genen verklaren, de erfelijkheid, kan gelijk zijn. De informatie over kwantitatieve sekseverschillen komt uit de test of de parameters g, c en e gelijk zijn voor jongens en meisjes.

RESULTATEN

Wanneer we kijken naar de gemiddelde scores op NoG (zie Tabel 5.1), valt op dat jongens gemiddeld hoger scoren dan meisjes. De tweelingcorrelaties (Tabel 5.4) geven een eerste indicatie van de relatieve invloed van genen en omgeving op NoG. De correlaties tussen MZ tweelingen zijn, bij zowel jongens als meisjes, bijna twee keer zo hoog als bij DZ tweelingen. Dit suggereert dat genetische factoren een rol spelen bij het verklaren van individuele verschillen in NoG. Bovendien suggereert de bijna twee keer zo hoge correlatie bij MZ tweelingen als bij DZ tweelingen dat er weinig invloed is van gedeelde omgeving. De correlaties suggereren ook dat unieke omgevingsinvloeden een belangrijke rol spelen, aangezien de correlaties tussen MZ tweelingen aanzienlijk lager zijn dan 1. De correlatie tussen tweelingen van ongelijk geslacht suggereert dat er kwalitatieve sekseverschillen zijn in de etiologie van NoG, omdat de correlatie kleiner is dan de correlaties tussen DZ tweelingparen van gelijke sekse. Zowel voor jongens als voor meisjes gold dat ze gemiddeld hoger scoren op NoG naarmate ze ouder worden, $b_{leeftijd}$=0,14 voor jongens, en $b_{leeftijd}$=0,18 voor meisjes.

Tabel 5.4. Tweelingcorrelaties, gecontroleerd voor leeftijd, naar zygositeit en geslacht

Zygositeit	Jongens	Meisjes	Ongelijk geslacht
MZ	.58	.62	-
DZ	.30	.35	.22

Genetische analyses

De *fit* van verschillende modellen zijn vergeleken met likelihood-ratio tests (zie Tabel 5.5). Wanneer de invloed van genetische factoren, gedeelde omgeving en unieke omgeving gelijkgesteld worden voor jongens en meisjes (model GCE_{gsv}), treedt geen significant verlies in *goodness of fit* op in vergelijking met een model mét sekseverschillen (model GCEsv). Wanneer de invloed van de gedeelde omgeving gelijk wordt gesteld aan nul (model GE_{gsv}), treedt ook geen significant verlies aan *fit* op. Wanneer de invloed van genetische factoren gelijk wordt gesteld aan nul (model E_{gsv}), is dat wel het geval. Tot slot leidt ook het gelijkstellen van de genetische correlatie tussen DZ tweelingen van ongelijke sekse aan 0,5 (model GEgsv+rG=0,5) tot een significant verlies in *fit*. Dit duidt op een kwalitatief sekseverschil, waarbij verschillende genetische factoren bij jongens en meisjes een rol spelen in de etiologie van NoG. Het uiteindelijke model met de beste *fit* is het model zonder sekseverschillen in de grootte van g, c en e, zonder significante invloed van gedeelde omgeving en mét kwalitatieve sekseverschillen (model GE). Zie Tabel 5.6 en Afbeelding 5.4 voor de resultaten uit dit model.

Het best passende model laat dus zien dat bij zowel jongens als meisjes ongeveer 61 procent van de variantie wordt verklaard door genetische factoren en 39 procent door unieke omgevingsfactoren. De gedeelde omgeving verklaart in het best passende model bij zowel jongens als meisjes geen variantie. Dit betekent dat er niet of nauwelijks omgevingsinvloeden zijn die leiden tot gelijke uitkomsten in tweelingen. Tot slot is de genetische correlatie tussen verschillende seksen lager dan 0,5 (geschat op 0,36). Dit duidt op een kwalitatief sekseverschil, waarbij bij NoG van jongens en meisjes verschillende genetische factoren een rol spelen.

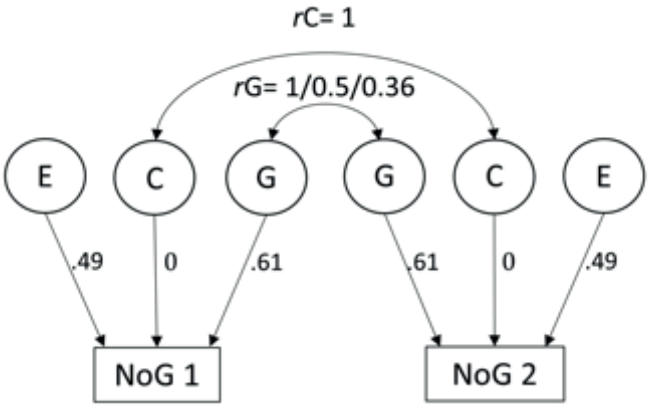
Tabel 5.5. Model fitting uitkomsten, en ruwe schattingen van de variantiecomponenten g^2, c^2 en e^2 met 95 procent betrouwbaarheidsintervallen

Model	Vergeleken met	-2 ll	AIC	df	Δ df	p		g^2 (95% BI)	c^2 (95% BI)	e^2 (95% BI)
GCE_{sv}	Gesatureerd	20.693,87	2.315,87	9.189	8	.679	♂	0,33 (0,23, 0,39)	0,02 (0, 0,12)	0,27 (0,24, 0,30)
	model						♀	0,33 (0,25, 0,41)	0,05 (0, 0,13)	0,23 (0,21, 0,25)
GCE_{gsv}	GCE_{sv}	20.700,16	2.316,16	9.192	3	.098	♂	0,34 (0,27, 0,39)	0,03 (0, 0,09)	0,24 (0,23, 0,26)
							♀	0,34 (0,27, 0,39)	0,03 (0, 0,09)	0,24 (0,23, 0,26)
GE_{gsv}	GCE_{sv}	20.701,20	2.315,20	9.193	4	.119	♂	0,37 (0,35, 0,39)	0	0,24 (0,23, 0,26)
							♀	0,37 (0,35, 0,39)	0	0,24 (0,23, 0,26)
E_{gsv}	GCE_{sv}	21.595,73	3.207,73	9.194	5	<.001	♂	0	0	0,61 (0,59, 0,63)
							♀	0	0	0,61 (0,59, 0,63)
$GE_{gsv+rG-0,5}$	GCE_{sv}	20.711,43	2.323,43	7.958	5	.004	♂	0,37 (0,34, 0,39)	0	0,25 (0,23, 0,26)
							♀	0,37 (0,34, 0,39)	0	0,25 (0,23, 0,26)

Tabel 5.6. Resultaten meest waarschijnlijke tweelingmodel (model 1)

Sekse	SA	SC	SE	b leeftijd	rG (m-v)
Jongens	,606	0	,394	0,14	0,36
Meisjes	,606	0	,394	0,18	0,36

Legenda: SA=gestandaardiseerde additieve genetische invloeden, SC=gestandaardiseerde gedeelde omgevingsinvloeden, SE=gestandaardiseerde unieke omgevingsinvloeden, b leeftijd=regressiecoëfficiënt NoG~leeftijd, rG (m-v)=genetische correlatie DZ tweelingen van ongelijke sekse.



Afbeelding 5.4. GEsv-model: Geeft de relatie tussen tweeling 1 en 2 weer, en de schattingen voor de paden g, c en e

Noot: rG=1 voor MZ tweelingparen, rG=0,5 voor DZ tweelingparen, en rG=0,36 voor tweelingparen van ongelijke sekse.

DISCUSSIE

Om na te gaan wat bij Nederlandse jongeren de intergenerationele continuïteit van NoG, zoals in meerdere onderzoeken gedocumenteerd, kan verklaren, deden we een onderzoek onder tweelingen. Met een tweelingendesign kan de relatieve invloed van genetische factoren, gedeelde omgeving en unieke omgeving op individuele verschillen in NoG worden geschat. We vonden een duidelijke bijdrage van genetische invloeden en geen bijdrage van gedeelde omgevingsfactoren. Hierbij impliceert de bijdrage van genetische invloeden genetische transmissie, en de afwezigheid van gedeelde omgevingsfactoren impliceert de afwezigheid van culturele transmissie van ouders op kinderen. Uit dit onderzoek volgt dat genetische transmissie belangrijk is, maar dat gedeelde omgevingsfactoren, en daarmee culturele transmissie, niet of amper een rol spelen op 13- tot 17-jarige leeftijd.

Dit onderzoek laat, in lijn met eerder onderzoek met vergelijkbare uitkomstmaten (Ferguson, 2010; Veroude e.a., 2016), zien dat genetische factoren een grote rol spelen bij NoG. De genetische factor representeert de invloed van een groot, waarschijnlijk zeer groot, aantal genetische varianten op NoG. De invloeden van deze varianten sommeren tot de totale genetische invloed. De resultaten uit dit onderzoek wijzen dus niet op het bestaan van één NoG- of criminaliteitsgen. Alleen bij een aantal zeldzame ziekten is er sprake van een directe relatie tussen een gen en een eigenschap, zoals bij de ziekte van Huntington (Donker, 2000).

Vaak ligt bij criminologisch onderzoek de nadruk vooral op omgevingsfactoren bij het verklaren van individuele verschillen en intergenerationele continuïteit van antisociaal en crimineel gedrag. In deze studie werd geen significante invloed van de gedeelde omgeving gevonden, en ook bij eerdere tweelingstudies met vergelijkbare uitkomstmaten waren de invloeden vaak klein (Ferguson, 2010; Veroude e.a., 2016). Dit impliceert dat culturele transmissie geen of slechts een kleine rol speelt in intergenerationele continuïteit van NoG. Het is daarom van belang dat in de criminologie rekening gehouden wordt met genetische factoren wanneer verschillen in NoG, of intergenerationele continuïteit van NoG, worden onderzocht. Dit betekent echter niet dat de omgeving geen invloed heeft, want unieke omgevingsinvloeden verklaren 40 procent van de variantie in NoG. Omgevingsfactoren leiden dus vooral tot verschillen in NoG voor kinderen uit eenzelfde gezin, en niet tot overeenkomsten. De bevindingen suggereren dat de nadruk bij het zoeken naar effectieve interventies vooral zou moeten liggen op de factoren die kindspecifiek zijn.

Ook bij eerder onderzoek werden vaak slechts kleine invloeden van de gedeelde omgeving gevonden. De schattingen lopen echter uiteen. Uit de meta-analyse van Ferguson (2010) volgde dat bij externaliserend probleemgedrag gemiddeld 11 procent van de variantie wordt verklaard door gedeelde omgevingsinvloeden. Een significant modererend effect van leeftijd op de invloed van de gedeelde omgeving werd niet gevonden. In een Nederlandse steekproef vonden Bartels en collega's (2005) dat gedeelde omgevingsfactoren 23 tot 36 procent van de variantie in NoG bij kinderen verklaren. De deelnemers in het onderzoek van Bartels e.a. waren kinderen tussen 7 en 12 jaar, terwijl de deelnemers in dit onderzoek tussen de 13 en 17 jaar waren. Uit eerder onderzoek met andere fenotypes, zoals IQ, sportgedrag of internaliserende problemen, is gebleken dat de invloed van gedeelde gezinsomgeving sterk kan afnemen met leeftijd (Bartels e.a., 2002; Huppertz e.a., 2016; Lamb e.a., 2010). Adolescenten gaan meer om met leeftijdsgenoten, zonder ouderlijke supervisie (Steinberg, 1986), en leeftijdsgenoten gaan mogelijk een belangrijkere rol spelen dan de ouders. Dit pleit ervoor dat eventuele interventies op gezinsniveau vooral moeten worden overwogen op jonge leeftijd. Overigens is een andere mogelijke verklaring het verschil in dataverzameling. In de studie van Bartels en collega's (2005) werd gebruik gemaakt van ouderrapportage, waarbij ouders informatie gaven over het gedrag van hun kinderen. In de huidige studie werd gebruik gemaakt van zelfrapportage, waarbij jongeren zelf vragen over hun gedrag beantwoorden. Bij onder andere angststoornissen en depressie is eerder gevonden dat de schattingen van gedeelde omgeving hoger waren wanneer gebruik werd gemaakt van ouderrapportage in vergelijking tot zelfrapportage (Lubke e.a., 2016). Omdat voor jonge kinderen zelfbeoordelingen lastig zijn, is het een uitdaging om de effecten van beoordelaar en leeftijd apart te bekijken.

In dit onderzoek vonden wij geen sterke aanwijzingen voor sekseverschillen in de relatieve invloed van genetische factoren en gedeelde of unieke omgeving. Door te kijken naar tweelingparen met ongelijk geslacht zijn wel kwalitatieve sekseverschillen gevonden. Uit de analyses, waaruit blijkt dat de genetische correlatie tussen tweelingen van ongelijke sekse lager is dan bij DZ tweelingparen van gelijke sekse, volgt dat verschillende genen mogelijk een rol spelen bij mannen en vrouwen.

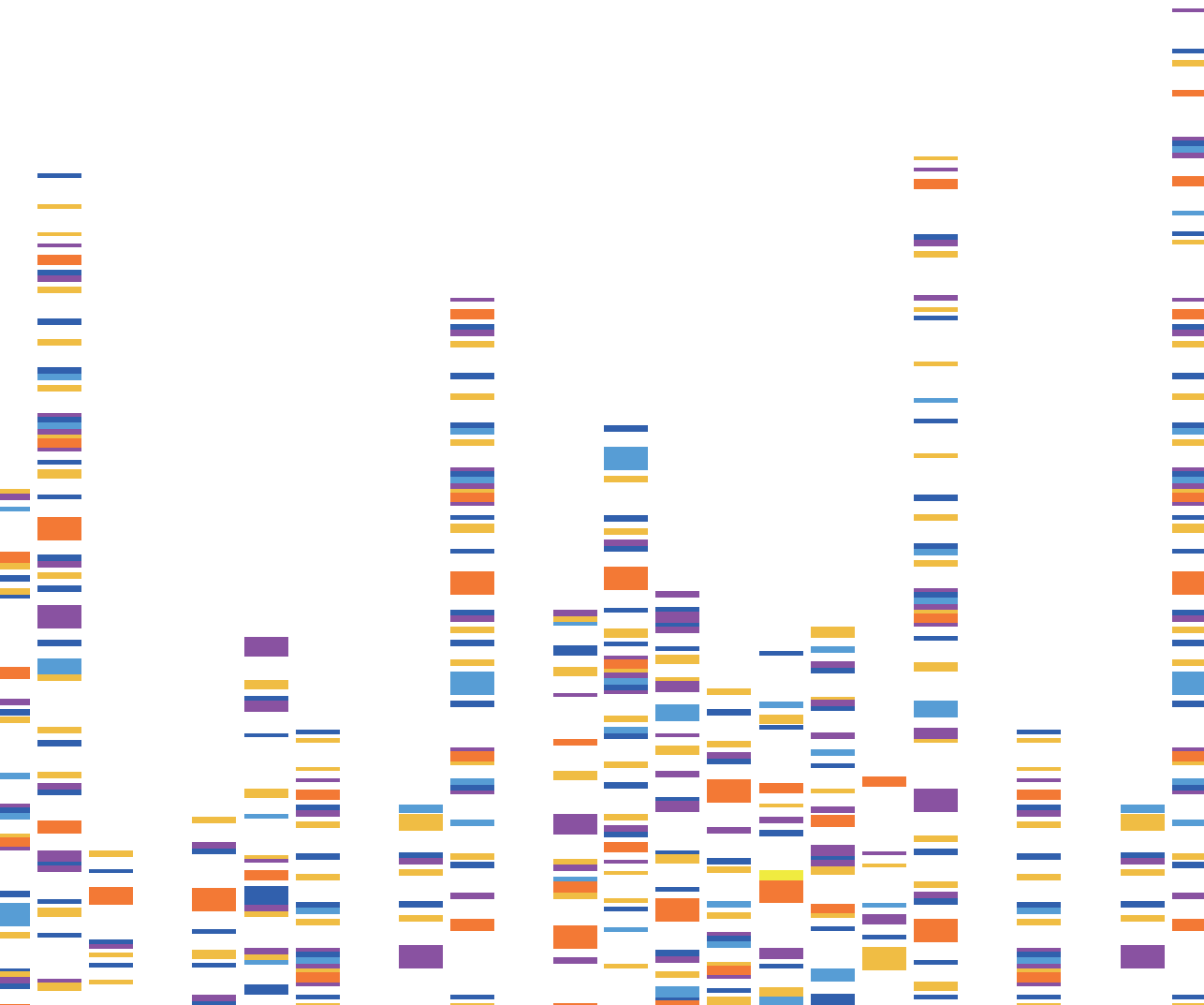
Tweelingonderzoek gaat gepaard met een aantal nuanceringen. Zo zijn er factoren die niet expliciet gemodelleerd worden in een klassiek tweelingmodel, die de schattingen kunnen beïnvloeden. In deze studie zijn niet-lineaire genetische effecten (interacties tussen genen of allelen) niet meegenomen

in het model. De belangrijkste reden daarvoor was overigens dat dergelijke effecten leiden tot een ander patroon van MZ en DZ correlaties dan werd gezien in deze studie. Genetische interactie-effecten resulteren in DZ correlaties die (veel) lager zijn dan de helft van de MZ correlatie. Ook kan, zoals besproken in de inleiding, gen-omgevingsinteractie of -correlatie een rol spelen. Wanneer de gedeelde omgeving kleiner zou zijn bij DZ tweelingen dan bij MZ tweelingen, kan dit leiden tot een overschatting van de erfelijkheid. Wanneer er 'assortative mating', waarbij ouders genetisch meer op elkaar lijken dan op basis van 'random mating', verwacht zou worden, leidt dit tot een overschatting van de DZ correlaties, terwijl de MZ correlaties gelijk blijven. Er vindt dan dus een overschatting van de gedeelde omgeving plaats. Deze implicaties van tweelingmodellen waren onderdeel van een discussie in *Criminology* (Burt & Simons, 2014; Barnes e.a., 2014), waarbij Burt en Simons openlijk twijfelden aan het nut en de validiteit van tweelingstudies binnen de criminologie. Zoals Barnes en collega's echter aangaven, zijn de gevolgen van schendingen van de aannames, en van aanwezigheid van niet-gemodelleerde effecten, vaak klein. Door tweelingmodellen uit te breiden met data van hun ouders, partners en/of kinderen kan de invloed van gen-omgevingscorrelatie, gen-omgevingsinteractie, dominantie en assortative mating verder onderzocht worden. Voor eigenschappen waarbij dit design is toegepast bij Nederlandse tweelingfamilies, zoals bijvoorbeeld attentieproblemen, borderline personality disorder, leesvaardigheid of alcoholgebruik, bleken resultaten goed overeen te komen met die op basis van het klassieke tweelingendesign (Boomsma e.a., 2010; Distel e.a., 2009; Swagerman e.a., 2017; Van Beek e.a., 2014).

De kracht van het huidige onderzoek ligt in de unieke onderzoeks-mogelijkheden die tweelingdata bieden. Binnen de criminologie, waar veel onderzoek primair naar omgeving kijkt, en mogelijke genetische 'confounding' negeert, is tweelingonderzoek van belang om een beeld te krijgen van de rol van zowel genen als omgeving bij het verklaren van verschillen in antisociaal en crimineel gedrag. Door genetische confounding kunnen de effecten van (vooral) factoren die tot de gedeelde omgeving behoren, overschat worden. Illustratief hiervoor is een onderzoek van Sariaslan en collega's (2013) onder kinderen in de drie grootste steden in Zweden (N=303.465), waarbij de invloed van buurteigenschappen op geweldsdelicten en drugsmisbruik wegvielen nadat gecontroleerd werd voor genetische factoren. Zo kunnen tweelingstudies, adoptiestudies of andere quasi-experimentele designs waarbij gebruik wordt gemaakt van genetische familiestructuren (D'Onofrio e.a., 2013), ook van belang zijn in onderzoek naar specifieke omgevingsfactoren.

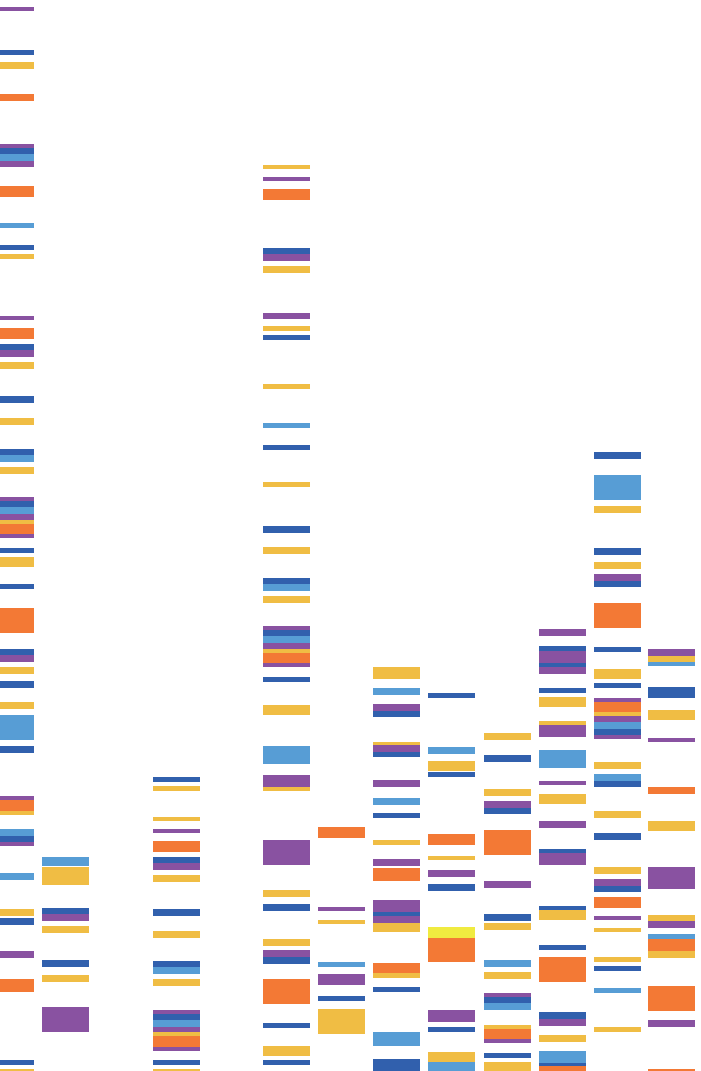
Dit onderzoek laat zien dat genetische factoren een belangrijke rol spelen in intergenerationele continuïteit. Over de specifieke genetische factoren die een rol spelen, weten we echter nog weinig tot niets. Een veelgebruikte methode om de relatie tussen specifieke genetische varianten en gedrag te testen, is een genoomwijde associatiestudie (GWAS). Hierbij worden associaties tussen – inmiddels soms ruim 7 miljoen – genetische varianten en gedrag onderzocht. Omdat het hier gaat om grote aantallen genetische varianten zijn ook zeer grote steekproeven nodig. Voor agressief, antisociaal en crimineel gedrag staat dit onderzoek in de kinderschoenen en zijn tot nu toe nog weinig significante associaties gevonden. Een tweetal voorbeelden van dergelijke studies zijn een onderzoek naar agressief gedrag (Pappa e.a., 2016) en naar antisociaal gedrag (Tielbeek e.a., 2017). Tielbeek en collega's vonden aanwijzingen dat verschillende genetische factoren mogelijk een rol spelen bij mannen en vrouwen, wat aansluit bij de kwalitatieve sekseverschillen die wij vonden in de huidige studie. Naast GWAS, wordt ook genoombreed epigenetisch onderzoek mogelijk. Er is een eerste onderzoek gedaan naar de invloed van DNA-methylatie op agressief gedrag bij volwassenen (Van Dongen e.a., 2015). Dit onderzoek liet zien dat epigenetische processen invloed kunnen hebben op agressief gedrag. Ook bij crimineel gedrag en NoG spelen epigenetische mechanismen mogelijk een rol.

De invloed van genetische en culturele transmissie is in deze studie onderzocht door te kijken naar variantie binnen één generatie 13- tot 17-jarigen. Dit betekent dat de schattingen een eerste indicatie zijn van de oorzaken van intergenerationele verbanden. In de toekomst zou meer onderzoek gedaan moeten worden waarbij meerdere generaties worden betrokken. Het ontbreken van een significante invloed van de gedeelde omgeving op de leeftijd 13 tot 17 jaar betekent dat een grote rol voor culturele transmissie niet waarschijnlijk is. De resultaten uit dit onderzoek zijn daarom een indicator dat, tijdens een levensfase waarin antisociaal gedrag zich ontwikkelt, intergenerationele continuïteit gedreven wordt door genetische transmissie, en dat culturele transmissie geen of slechts een kleine rol speelt.



6

Direct and Indirect Genetic Effects on Aggression



ABSTRACT

Background: Family members resemble each other in their propensity for aggression. A large proportion of this resemblance can be explained by genetic influences. If there are genotype-environment correlation mechanisms, however, genetic influences may in part reflect environmental influences. Such indirect genetic effects may arise, for example, by environmental manifestations of parental and sibling genotypes. In this study we investigate the importance of these 'indirect genetic effects' on aggression.
Method: We model the effect of polygenic scores based on three discovery genome wide association studies, namely early-life aggression, educational attainment, and ADHD. The associations with aggression were tested in a within- and a between-family design (N= 37,796 measures from 7,740 individuals, from 3,107 families, 55% female) and in a transmitted/non-transmitted PGSs design (42,649 measures from 6,653 individuals, from 3,024 families, 55% female).
Results: No evidence for contributions of indirect genetic effects on aggression were found, neither in a within-/ between-family design nor in a transmitted/non-transmitted PGSs design. Results indicate significant direct genetic effects on aggression for the PGSs based on early-life aggression, educational attainment, and ADHD.
Conclusion: The main drivers of familial clustering and intergenerational transmission of aggression are direct genetic influences, without evidence for genotype-environment correlation.

Based on:
van der Laan, C.M., van de Weijer, S.G.A., Pool, R., Hottenga, J., Nivard, M.G., & Boomsma, D.I. (submitted for publication). Direct and indirect genetic effects on aggression.

DIRECT AND INDIRECT GENETIC EFFECTS ON AGGRESSION

Aggression is behavior that aims to cause harm to others (Berkowitz, 1993; Lorenz, 1966). Aggression is relatively common, but the propensity for such behavior varies between individuals. It is well established that family members resemble each other in their propensity for aggression (Frisell, Lichtenstein, & Långström, 2011; Margolin et al., 2016, Repetti et al., 2002; van de Weijer et al., 2014; Veroude et al., 2016; Yu & Gamble, 2008; Van der Laan et al., 2021-b). This resemblance is the result of family influences, related to the psychosocial environment, social economic circumstances, and genes (Tolan, Dodge, & Rutter, 2013; Labella & Masten, 2018). Because family members share both their environment and their genes, it is difficult to disentangle effects of environment and genes (Plomin et al., 1977), unless the question is addressed in a study design that includes family members with different degrees of genetic relatedness, such as adoption or twin studies.

Twin studies suggest that clustering of aggression in families is predominantly due to genetic effects, as around 50% the variance in aggression can be explained by genetic variance, and only a very small amount by shared environmental influences, i.e., environmental influences that lead to phenotypical similarities between twins (Odintsova et al., 2019; Veroude et al., 2016). However, twin and adoption studies come with assumptions, including that genetic and environmental effects are uncorrelated. The environment in which parents raise their children, is partly dependent on the parental genotypes. Because offspring inherit their genotype from their parents and are exposed to the rearing environment created by their parents, a correlation between the genotype and environment can be induced (Plomin, DeFries & Fulker, 1977; Kendler & Eaves, 1986; Plomin, 2014). The rearing environment will depend in part on the parental genotypes and if information is available on parental and offspring genotypes, it is possible to distinguish between the direct effects of genetic transmission from parents to offspring and the indirect genetic effects through the rearing environment (Bates et al., 2018; Kong et al., 2018).

In this study we employ two polygenic scores (PGS) to separate direct and indirect genetic effects. The PGS are based on results from genome-wide association studies (GWAS) of early-life aggression (EL-AGG), educational attainment (EA) and Attention-Deficit Hyperactivity Disorder (ADHD).

We selected these discovery studies because of their sample sizes and the genetic correlations of these traits with aggression (Ip et al., 2021). The first (transmitted) PGS reflects the inheritance of genetic variants that directly increases the propensity for a certain trait, while the second (non-transmitted) PGS reflects the indirect genetic effects, i.e., the effects of genetic variants that were not transmitted from parents to offspring but which may influence the outcome in offspring through environmental influences. We expect non-transmitted genetic effects could play a role in aggression, because this has previously been demonstrated for associated traits, such as EA (Kong et al., 2018; Demange et al., 2020).

Polygenic scores

PGSs are the sum of trait-associated alleles (coded by 0, 1, or 2 for the presence of a (risk) increasing allele) across the genome, weighted by their effect size. In less than two decades, GWAS have identified robust associations between tens of thousands of single base pair variants (single-nucleotide polymorphisms [SNPs]) and complex phenotypes (Visscher et al., 2012; Visscher et al., 2017). Individual SNP effects on most human complex traits tend to be small, but as the power of GWAS grows (Mills & Rahal, 2020), so does our ability to combine the effect of multiple SNPs and construct genome-wide polygenic scores related to complex traits. As such, PGSs are dependent on well powered high quality GWAS results.

Within-family analysis

An alternative way to separate direct and indirect genetic effects is to use a within- and between-family design. Whereas the transmitted/ non-transmitted PGS design depends on the availability of genotyped parent-offspring families, a within- and between-family design depends on the availability of genotyped siblings and dizygotic twins. Siblings and dizygotic twins share on average 50 percent of their genotypes. Quantitative genetic theory predicts that the expected proportion of alleles shared identically by descent (IBD) between siblings has a mean of 0.5 and a standard deviation (SD) of 0.039 (Visscher et al., 2006). Hemani et al. (2013) confirmed these predictions by analyzing ~20,000 SNPs in linkage equilibrium for 20,240 sibling pairs and obtained a SD of 0.037. Because the transmission of alleles from parents to offspring is randomized during meiosis, siblings have an equal probability of inheriting any given allele from father and mother, so that the expected correlations between PGSs in siblings are equal to the expected proportion of alleles shared IBD, i.e., r= 0.5. By modeling the effect

of PGSs on aggression in a within family design, we can test for an association between the PGSs and aggression while eliminating potential confounding by family-environment (e.g., Selzam et al., 2019; Plomin, 2014).

The current study

In this study we investigate whether direct and indirect genetic effects play a role in the familial clustering of aggression. We test the hypotheses that a) there is an association between the transmitted PGS and aggression at the population level (also called between-family PGS prediction) and that b) there is an association between the transmitted PGS and aggression within families. If a difference in the strength of the association is observed, this may indicate that indirect genetic effects drive part of the between-family genetic effects. Secondly, we test the hypothesis that non-transmitted PGSs influence individual differences in aggression.

For our analysis we analyze PGSs based on summary statistics of three large GWASs. First, we include early-life aggression (EL-AGG), the largest GWAS on aggression to date, with an effective sample size of N= 151,741 (Ip et al., 2021). In a previous study, it was demonstrated that a PGS of EL-AGG also predicts aggression in adults (van der Laan et al., 2021-c). Secondly, we include a PGS of Attention-Deficit Hyperactivity Disorder (ADHD; Demontis et al., 2019) because Ip and colleagues (2021) found a perfect genetic correlation (rg) with childhood aggression, rg= 1.00 (se= 0.07), and because the effective sample size and SNP heritability, i.e., the variance in ADHD explained by SNP-variance, are much larger than in the EL-AGG GWAS, indicating more statistical power. Third, we include a PGS of educational attainment (EA; Lee et al., 2018) because aggression is associated with academic performance (Vuoksimaa et al., 2021) and the discovery GWAS is among the most highly powered GWASs of a behavioral phenotype to date. Based on the genetic association analyses of EL-AGG and EA, a genetic correlation of rg= -0.50 (se= 0.04) was estimated (Ip et al., 2021).

METHODS

Participants

All participants were enrolled in the Netherlands Twin Register (NTR; Ligthart et al., 2019). Consent was obtained from all participants, or, for children, from their parents. All participants are twins, or family members of twins. For siblings of young twins, phenotyping mainly comes from teacher

reports. The teacher report, parent report, and age specific self-report data collections were structured so that twins and siblings were recruited based on the age of the twin. For this study, data-collection waves were restructured temporarily, so that siblings of twins were scored relative to participants of similar age, with Item-Response Theory models (IRT). See Table 6.1 for an overview of the data-collection waves, with phenotyping sample sizes, ages, and IRT aggression scores.

For the within- and between- family PGS prediction, all genotyped participants with at least one DZ twin or sibling in the same data collection wave were included in the analyses (37,796 measures from 7,740 individuals, from 3,107 families). See Table 6.2 for an overview of the samples per rater and data-collection wave. For the transmitted and non-transmitted PGS prediction, genotyped participants with two genotyped parents were included in the analyses (42,649 measures from 6,653 individuals, from 3,024 families). See Table 6.3 for an overview of the samples per rater and data-collection wave.

Phenotyping

Participants were phenotyped by parental ratings, starting at age 3 years, by teacher report, starting at age 5, and/or by self-report, starting at age 12. Teachers completed the aggression scale from the ASEBA 6-18 years Teacher Report Form (TRF, 20 items; Achenbach & Rescorla, 2001). Parents completed the aggression scale from the age 1.5-5 years Child Behavior Checklist (CBCL, 19 items; Achenbach & Rescorla, 2000) or age 6-18 years Child Behavior Checklist (CBCL, 18 items; Achenbach & Rescorla, 2001) and/or Devereux Child Behavior rating scale (DCB, 7 items; Spivack and Spotts, 1966; van Beijsterveldt et al., 2004). Self-reports were obtained with the ASEBA Youth Self Report (YSR, 17 items; Achenbach & Rescorla, 2001) and/or ASEBA Adult Self Report (ASR, 15 items; Achenbach & Rescorla, 2003). All items on the TRF, CBCL, YSR, YASR, and ASR were scored on a three-level scale: 0= never, 1= sometimes, 2= often. Items on the DCB were scored on a five-level scale: 0= never, 5= frequently. CBCL and TRF data were collected as a function of child age, and YSR and ASR were collected in adolescent and adult surveys in specific timeframes.

Aggression Item-Response Theory scores (IRTagg) were calculated regardless of genotyping status, with the Generalized Partial Credit Model (GPCM) in R (R Core Team, 2017), with the mirt package (Chalmers, 2012). GPCM is a form of Item-Response Theory (IRT), developed specifically to

analyze polytomous data (Embretson & Reise, 2000). IRT appropriately weights the relative contributions of the individual aggression items, resulting in a scale with a more favorable distribution than a sum-score, and can handle missing item data. Aggression scores were calculated for all waves of NTR data collection separately (e.g., for all 7-year-olds or for all participants in the survey collected in 1993). By fitting a separate model for each wave of data collection, aggression scores for each participant are relative to all other participants in that wave of data collection, thereby filtering out potential 'wave' or data collection effects. Because the IRT score for each individual is relative to all other participants in the same wave of data collection, the mean score for each wave is zero. All participants with a maximum of 20 percent missing items on the different aggression scales were included in the GPCMs.

Table 6.1. Descriptive statistics phenotyping sample: N, Age, and Aggression by sex for each instrument

	N obs	N pers	N fam	Mean Age (SD)	Mean IRTagg (SD)	N male	Mean Age Males (SD)		Mean IRTagg Males (SD)		N female	Mean Age Females (SD)		Mean IRTagg Females (SD)	
TRF age 5	1,132	1,132	641	5.58 (0.51)	0 (0.84)	561	5.57	(0.52)	0.17	(0.90)	571	5.60	(0.50)	-0.16	(0.74)
TRF age 7	8,891	8,890	4,946	6.96 (0.23)	0 (0.83)	4,386	6.96	(0.23)	0.17	(0.90)	4,505	6.96	(0.22)	-0.17	(0.71)
TRF age 10	20,412	18,187	9,630	9.07 (0.86)	0 (0.84)	10,155	9.06	(0.86)	0.18	(0.90)	10,252	9.09	(0.86)	-0.17	(0.72)
TRF age 12	14,809	13,966	7,812	12.00 (0.63)	0 (0.83)	7,347	12.01	(0.63)	0.17	(0.90)	7,455	11.99	(0.63)	-0.17	(0.71)
CBCL V age 3	26,313	26,313	13,197	3.38 (0.50)	0 (0.94)	13,152	3.38	(0.50)	0.09	(0.95)	13,158	3.37	(0.49)	-0.09	(0.91)
DCB V age 5	31,771	31,771	15,832	5.49 (0.66)	0 (0.87)	15,725	5.49	(0.65)	0.10	(0.89)	16,040	5.48	(0.66)	-0.10	(0.84)
CBCL V age 7	17,883	17,883	8,941	7.41 (0.49)	0 (0.90)	8,950	7.42	(0.49)	0.11	(0.93)	8,927	7.40	(0.49)	-0.11	(0.86)
CBCL V age 10	16,215	15,674	7,843	9.96 (0.59)	0 (0.89)	8,019	9.96	(0.60)	0.10	(0.93)	8,192	9.95	(0.59)	-0.10	(0.85)
CBCL V age 12	12,341	12,299	6,166	12.28 (0.48)	0 (0.88)	6,047	12.28	(0.48)	0.08	(0.92)	6,290	12.29	(0.48)	-0.07	(0.84)
CBCL M age 3	38,877	38,877	19,451	3.36 (0.49)	0 (0.94)	19,376	3.36	(0.49)	0.11	(0.96)	19,493	3.36	(0.49)	-0.11	(0.91)
DCB M age 5	35,666	35,666	17,779	5.49 (0.66)	0 (0.86)	17,687	5.49	(0.65)	0.11	(0.90)	17,973	5.49	(0.67)	-0.11	(0.82)
CBCL M age 7	25,306	25,306	12,661	7.43 (0.50)	0 (0.91)	12,623	7.43	(0.50)	0.12	(0.94)	12,675	7.42	(0.49)	-0.12	(0.87)
CBCL M age 10	23,357	22,586	11,309	9.95 (0.62)	0 (0.91)	11,560	9.95	(0.63)	0.11	(0.94)	11,792	9.95	(0.62)	-0.11	(0.86)
CBCL M age 12	17,535	17,457	8,735	12.29 (0.47)	0 (0.90)	8,662	12.28	(0.47)	0.08	(0.93)	8,868	12.30	(0.48)	-0.08	(0.85)
YSR 1991	3,327	3,327	1,668	17.95 (2.24)	0 (0.86)	1,501	17.93	(2.25)	-0.05	(0.90)	1,826	17.97	(2.24)	0.04	(0.82)
YSR 1995	3,344	3,344	1,708	19.98 (3.10)	0 (0.85)	1,477	19.96	(3.09)	-0.06	(0.88)	1,866	19.99	(3.11)	0.05	(0.83)
YSR 1997	4,715	4,715	1,999	26.73 (10.46)	0 (0.85)	1,903	26.27	(10.52)	-0.04	(0.86)	2,812	27.05	(10.41)	0.03	(0.85)
YSR 2000	6,702	6,702	3,172	30.49 (10.78)	0 (0.85)	2,514	30.3	(10.45)	-0.06	(0.85)	4,183	30.6	(10.98)	0.04	(0.85)
ASR 2009	15,048	15,048	6,805	41.49 (15.40)	0 (0.85)	5,353	43.53	(15.89)	-0.11	(0.80)	9,693	40.36	(15.00)	0.06	(0.88)
ASR 2014	16,203	16,203	7,630	40.19 (14.62)	0 (0.85)	5,838	42.04	(15.09)	-0.11	(0.79)	10,362	39.15	(14.24)	0.06	(0.88)
YSR age 14	8,672	8,551	4,791	14.64 (0.59)	0 (0.87)	3,718	14.65	(0.60)	0.04	(0.91)	4,951	14.63	(0.59)	-0.03	(0.84)
YSR age 16	7,898	7,293	4,319	16.70 (0.46)	0 (0.87)	3,308	16.7	(0.46)	0.01	(0.91)	4,587	16.7	(0.46)	-0.01	(0.83)
YSR age 18	4,384	4,099	2,594	18.74 (0.91)	0 (0.84)	1,675	18.70	(0.90)	0.02	(0.89)	2,708	18.76	(0.91)	-0.01	(0.82)
Total	360,801	83,027	33,653	11.57 (11.50)	0 (0.89)	171,541	10.75	(10.72)	0.09	(0.92)	189,180	12.32	(12.11)	-0.08	(0.85)

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = Youth Self-Report, ASR = Adult Self-report, N obs= number of observations, N pers= Number of participants, N fam= number of families, SD= standard deviation, IRTagg= Item-Response Theory aggression score.
Note: The measure reflects the measurement wave used for phenotyping. The family structure in the data means that originally, siblings of twins were measured with the same instrument as the twins. Because some measurement waves are age dependent, siblings were restructured into the correct age-dependent phenotyping samples. For 80 measures sex was missing, or the participant changed gender.

Table 6.2. Descriptive statistics between- & within-family sample: N, Age, and Aggression by sex for each instrument

Measure	N pers	N fam	Mean Age (SD)		Mean IRTagg (SD)		N male	Mean Age Males (SD)		Mean IRTagg Males (SD)		N female	Mean Age Females (SD)		Mean IRTagg Females (SD)	
TRF age 5	140	70	5.79	(0.41)	-0.09	(0.79)	70	5.74	(0.44)	0.09	(0.88)	70	5.83	(0.38)	-0.27	(0.66)
TRF age 7	1,384	609	7.75	(1.10)	-0.05	(0.80)	692	7.76	(1.13)	0.12	(0.86)	692	7.74	(1.07)	-0.23	(0.68)
TRF age 10	1,492	685	9.92	(0.89)	0	(0.83)	721	9.90	(0.88)	0.16	(0.91)	771	9.94	(0.91)	-0.15	(0.72)
TRF age 12	1,091	529	12.08	(0.79)	-0.01	(0.84)	518	12.10	(0.74)	0.19	(0.92)	573	12.07	(0.83)	-0.19	(0.72)
CBCL V age 3	1,890	937	3.38	(0.49)	-0.06	(0.96)	914	3.40	(0.50)	0.00	(0.98)	976	3.36	(0.49)	-0.11	(0.94)
DCB V age 5	2,384	1,183	5.59	(0.75)	0.06	(0.88)	1,136	5.62	(0.77)	0.15	(0.89)	1,248	5.56	(0.73)	-0.02	(0.86)
CBCL V age 7	1,937	963	7.43	(0.57)	0.02	(0.90)	924	7.43	(0.57)	0.12	(0.93)	1,013	7.42	(0.57)	-0.08	(0.87)
CBCL V age 10	1,956	972	9.91	(0.68)	-0.01	(0.89)	940	9.91	(0.68)	0.05	(0.93)	1,016	9.90	(0.68)	-0.06	(0.85)
CBCL V age 12	1,416	703	12.16	(0.45)	0.01	(0.89)	660	12.15	(0.44)	0.08	(0.95)	756	12.17	(0.47)	-0.05	(0.82)
CBCL M age 3	2,524	1,253	3.35	(0.49)	-0.05	(0.98)	1,215	3.37	(0.49)	0.04	(0.98)	1,309	3.33	(0.48)	-0.14	(0.97)
DCB M age 5	2,524	1,252	5.59	(0.74)	0.04	(0.87)	1,211	5.63	(0.76)	0.13	(0.90)	1,313	5.56	(0.72)	-0.05	(0.84)
CBCL M age 7	2,343	1,165	7.43	(0.56)	0	(0.93)	1,128	7.43	(0.57)	0.10	(0.97)	1,215	7.42	(0.56)	-0.09	(0.89)
CBCL M age 10	2,461	1,222	9.92	(0.70)	-0.02	(0.92)	1,178	9.93	(0.71)	0.04	(0.96)	1,283	9.91	(0.70)	-0.08	(0.87)
CBCL M age 12	1,710	850	12.18	(0.46)	0	(0.91)	797	12.16	(0.45)	0.05	(0.94)	913	12.20	(0.47)	-0.05	(0.87)
YSR 1991	583	291	17.74	(2.28)	0.02	(0.90)	252	17.62	(2.27)	-0.04	(0.95)	331	17.83	(2.28)	0.06	(0.86)
YSR 1995	680	339	20.04	(3.17)	0	(0.84)	280	20.08	(3.17)	-0.04	(0.86)	400	20.02	(3.18)	0.03	(0.83)
YSR 1997	1,788	649	28.19	(11.69)	-0.03	(0.86)	722	27.80	(11.94)	-0.07	(0.85)	1,066	28.45	(11.52)	0.00	(0.86)
YSR 2000	1,921	729	32.48	(11.85)	-0.04	(0.84)	693	32.12	(12.51)	-0.10	(0.84)	1,228	32.69	(11.46)	-0.01	(0.84)
ASR 2009	2,193	900	36.61	(13.76)	0.04	(0.86)	755	36.65	(14.33)	-0.09	(0.81)	1,438	36.59	(13.46)	0.10	(0.88)
ASR 2014	1,985	843	34.30	(12.13)	0.02	(0.86)	676	32.97	(11.92)	-0.08	(0.79)	1,309	34.99	(12.19)	0.07	(0.88)
YSR age 14	1,345	585	15.32	(1.21)	-0.03	(0.83)	569	15.23	(1.12)	-0.01	(0.87)	776	15.39	(1.27)	-0.05	(0.80)
YSR age 16	1,357	583	17.26	(1.23)	-0.1	(0.82)	569	17.17	(1.18)	-0.09	(0.87)	788	17.33	(1.25)	-0.11	(0.79)
YSR age 18	464	191	18.28	(1.48)	0.17	(0.84)	201	18.28	(1.51)	0.25	(0.87)	263	18.29	(1.46)	0.10	(0.81)
YSR Pilot	228	99	16.49	(1.19)	0.31	(0.85)	102	16.43	(1.27)	0.43	(0.97)	126	16.53	(1.11)	0.21	(0.72)
Total	7,740	3,107	14.20	(12.03	0	(0.88)	16,923	12.98	(11.02)	0.05	(0.92)	20,873	15.20	(12.70)	-0.05	(0.86)

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = Youth Self-Report, ASR = Adult Self-report, N pers= Number of participants, N fam= number of families, SD= standard deviation, IRTagg= Item-Response Theory aggression score.

Note: The measure reflects the measurement wave for twins. Siblings of twins are included in the same measurement wave, so that the family structure in the data is preserved. Because some measurement waves are age dependent, some siblings were phenotyped as part of other, correct age-appropriate instrument samples. The total number of measures was 37,796, from 7740 individuals.

Table 6.3. Descriptive statistics transmitted/non-transmitted sample: N, Age, and Aggression by sex for each data collection wave

	N pers	N obs	N fam	Mean Age (SD)		Mean IRTagg (SD)		N male	Mean Age Males (SD)		Mean IRTagg Males (SD)		N female	Mean AGE Females (SD)		Mean IRTagg Females (SD)	
TRF age 5	122	122	66	5.64	(0.48)	0	(0.82)	57	5.56	(0.50)	0	(0.84)	65	5.71	(0.46)	0	(0.82)
TRF age 7	1,687	1,687	849	7.68	(1.04)	-0.07	(0.78)	847	7.67	(1.04)	0.08	(0.85)	840	7.70	(1.04)	-0.22	(0.67)
TRF age 10	1,886	1,886	996	9.84	(0.87)	-0.05	(0.81)	879	9.82	(0.88)	0.14	(0.89)	1,007	9.86	(0.86)	-0.21	(0.69)
TRF age 12	1,398	1,398	776	12.15	(0.67)	0	(0.83)	669	12.16	(0.67)	0.23	(0.90)	729	12.15	(0.67)	-0.20	(0.69)
CBCL V age 3	2,556	2,556	1324	3.46	(0.52)	-0.09	(0.92)	1,232	3.48	(0.54)	-0.06	(0.92)	1,324	3.45	(0.51)	-0.12	(0.92)
DCB V age 5	3,238	3,238	1665	5.61	(0.74)	-0.06	(0.87)	1,516	5.66	(0.75)	0.03	(0.89)	1,722	5.58	(0.72)	-0.14	(0.85)
CBCL V age 7	2,343	2,343	1219	7.46	(0.60)	0	(0.91)	1,110	7.48	(0.61)	0.10	(0.93)	1,233	7.44	(0.60)	-0.09	(0.88)
CBCL V age 10	2,499	2,499	1297	9.87	(0.72)	-0.04	(0.89)	1,168	9.88	(0.73)	0.03	(0.93)	1,331	9.85	(0.72)	-0.10	(0.85)
CBCL V age 12	1,519	1,519	791	12.17	(0.46)	0.02	(0.89)	659	12.18	(0.45)	0.08	(0.94)	860	12.16	(0.47)	-0.03	(0.84)
CBCL M age 3	3,347	3,347	1734	3.40	(0.51)	-0.09	(0.97)	1,588	3.42	(0.53)	-0.01	(0.97)	1,759	3.38	(0.50)	-0.16	(0.97)
DCB M age 5	3,353	3,353	1726	5.62	(0.74)	-0.04	(0.87)	1,583	5.66	(0.75)	0.07	(0.89)	1,770	5.58	(0.72)	-0.14	(0.84)
CBCL M age 7	2,788	2,788	1446	7.46	(0.60)	-0.05	(0.92)	1,332	7.48	(0.60)	0.03	(0.94)	1,456	7.45	0.60)	-0.12	(0.89)
CBCL M age 10	3,007	3,007	1556	9.87	(0.74)	-0.07	(0.91)	1,423	9.88	(0.73)	-0.01	(0.94)	1,584	9.87	(0.75)	-0.13	(0.88)
CBCL M age 12	1,770	1,770	921	12.19	(0.47)	0	(0.91)	792	12.19	(0.46)	0.05	(0.95)	978	12.19	(0.48)	-0.04	(0.87)
YSR 1991	755	755	416	17.65	(2.23)	0.01	(0.86)	303	17.68	(2.21)	-0.01	(0.93)	452	17.63	(2.24)	0.02	(0.82)
YSR 1995	972	972	544	19.95	(3.06)	-0.01	(0.84)	384	19.95	(3.07)	-0.06	(0.86)	588	19.95	(3.05)	0.03	(0.83)
YSR 1997	1,269	1,269	572	22.73	(4.86)	0.01	(0.86)	533	22.52	(4.63)	-0.03	(0.86)	736	22.89	(5.02)	0.03	(0.86)
YSR 2000	1,329	1,329	666	25.78	(5.25)	-0.01	(0.84)	504	25.48	(4.74)	-0.04	(0.81)	825	25.96	(5.53)	0.01	(0.86)
ASR 2009	1,967	1,967	1132	29.96	(8.65)	0.08	(0.88)	635	29.24	(7.91)	-0.05	(0.81)	1,332	30.30	(8.97)	0.15	(0.90)
ASR 2014	1,977	1,977	1165	31.40	(9.10)	0.04	(0.85)	661	30.45	(8.66)	-0.08	(0.76)	1,316	31.88	(9.28)	0.10	(0.89)
YSR age 14	1,117	1,117	556	15.21	(1.07)	0.03	(0.84)	446	15.15	(0.92)	0.04	(0.85)	671	15.25	(1.16)	0.02	(0.83)
YSR age 16	1,186	1,186	595	17.32	(1.09)	-0.04	(0.83)	491	17.24	(1.05)	-0.03	(0.91)	695	17.37	(1.11)	-0.05	(0.77)
YSR age 18	415	415	185	18.31	(1.36)	0.11	(0.82)	194	18.31	(1.36)	0.19	(0.87)	221	18.30	(1.35)	0.04	(0.77)
YSR Pilot	147	147	67	16.52	(1.22)	0.25	(0.84)	68	16.51	(1.40)	0.29	(0.97)	79	16.53	(1.06)	0.21	(0.71)
Total	6,652	42,647	3024	11.81	(8.68)	-.03	(0.88)	19,074	10.95	(7.75)	0.03	(0.91)	23,573	12.50	(9.31)	-0.08	(0.86)

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = ASEBA Youth Self-Report, ASR = ASEBA Adult Self-report, N pers= Number of participants, N obs= number of observations, N fam= number of families, SD= standard deviation, IRTagg= Item-Response Theory aggression score.
Note: The measure reflects the measurement wave for twins. Siblings of twins are included in the same measurement wave, so that the family structure in the data is preserved. Because some measurement waves are age dependent, some siblings were phenotyped as part of other, correct age-appropriate instrument samples.

Genotype data

Participants were genotyped as described in Chapter 4.

> Participants were genotyped on multiple platforms: Affymetrix Axiom, Affymetrix 6.0, Illumina 1M, Illumina 660, Illumina GSA, Perlegen Affymetrix. Samples with call rate <0.90, Plink heterozygosity F< −0.10 or F> 0.10, and inconsistency of X chromosome genotypes with reported gender were excluded. SNPs with MAF <1.0E-6, HWE p-value < 1.0E-6, and/or call rate <0.95 were removed. Genotype data were aligned with the 1000 Genomes reference panel, and filtered for SNPs with allele frequency differences from the CEU population larger than 0.20, palindromic SNPs, and DNA strand issues. DNA Identity By Descent (IBD) state was estimated for all individual pairs using Plink (Purcell et al., 2007) and King (Manichaikul et al., 2010) based on ~10.8k SNPs that all platforms have in common. Samples were removed if IBD did not match expected family relations. CEU population outliers were removed from the data with Smartpca software, based on per platform 1000 Genomes PC projection. Per platform, data were phased using Eagle and imputed to 1000 Genomes with Minimac (Das et al., 2016). The final merged genotype data consist of 12,152,830 SNPs. (p. 62)

Polygenic score construction

We obtained GWAS summary statistics from the Ip et al. (2021) early-life aggression GWAMA, Demontis et al. (2019) ADHD GWAMA, and Lee et al. (2018) educational attainment GWAMA, after leaving out all participants from the NTR. The GWAS effect sizes were used to calculate polygenic scores using SBayesR V2.03 (Lloyd-Jones et al., 2019) with default settings. For the between- and within-family subsamples, PGSs were calculated for each individual. In the within- and between-family analyses, the between-family element is the average of all family member's $\overline{PGS}_j$, the within-family element is the deviation of individual family members from that average ($\overline{PGS}_{ij}$-PGS_j). For the transmitted/non-transmitted subsamples, two PGSs were computed for each individual: the transmitted PGS (T-PGS) and the non-transmitted PGS (NT-PGS).

Analyses

Within/between family PGS prediction

First, we modeled the within- and between-family effects between the PGSs and aggression for each measurement wave (*N*= 24), in a mixed effects model with the package lme4 (Bates, 2015) in R (R core team, 2017). The model includes two fixed effects to separate the total effect between the transmitted PGSs and aggression into within- and between-family effects:

$$AGG_{ij} = \alpha_{0j} + \beta W * (PGS_{ij} - \overline{PGS_j}) + \beta_B * \overline{PGS_j} + \beta_{3\text{-}v} * x_{3\text{-}v} + \varepsilon_{ij} \quad \text{Equation 1.}$$

AGG denotes the IRT aggression outcome, *i* are the individual twins or siblings that are clustered within family *j*, so that PGSij is the polygenic score of individual *i* in family *j*, and $\overline{PGS_j}$ is the mean PGS value in family *j*. PGS*ij*–$\overline{PGS_j}$ gives the individual deviation of individual *i* from the family average ($\overline{PGS_j}$). The notation α_{0j} represents the overall intercept and deviation from that intercept in family *j*, ε_{ij} denotes the independent random error for individual *i* in family *j*. The between-family fixed effect β_B represents the expected change in aggression given a one standard deviation (SD) change in the family GPS average, and the within-family effect β_W represents the expected change given a one SD change in the difference between the individual PGS and the family average PGS. By including both β_W and β_B in the same model, the individual estimates are adjusted for, and independent of, the effect of the other estimate. $\beta_{3\text{-}v}$ represent the expected change in the aggression outcome given a one SD change in the included covariates $x_{3\text{-}v}$ (age, sex, dummy variables for genotyping arrays, and five ancestry-based principal components). The random intercept accounts for any additional genetically or environmentally invoked dependence between measures in twins or siblings, unaccounted for by the mean family PGS.

Second, the results from the 24 data collection waves were meta-analyzed with a fixed-effects model in Metafor (Viechtbauer, 2010) in R. Because individuals may be phenotyped multiple times and by multiple raters, the results from the 24 data collection waves are not independent. We quantify the dependence between subsamples by calculating the elements in the variance-covariance matrix of the sampling errors (*V*):

$$V_{kl} = se_k\left(\frac{N_{kl} r_p}{\sqrt{N_k N_l}}\right) se_l \quad \text{Equation 2.}$$

Where *V*kl is the element in the variance-covariance matrix *V* of the sampling errors for subsamples *k* and *l*, *se*k is the standard error for the estimate is subsample *k*, *se*l is standard error of the estimated effect in subsample *l*, *N*k*l* is the overlap between subsamples *k* and *l*, *N*k and *N*l are the sample sizes for subsamples *k* and *l*. The meta-analysis model is then given by , where *y* is a vector with the observed subsample outcomes, θ is the (average) true outcome, and *V* is the variance-covariance matrix of the sampling errors. See Supplements Table S6.I and S6.2 for the sample overlap matrix and cross-sample correlation matrix. Post-hoc tests of between-rater contrasts were performed by meta-analyzing the results with *rater* included as moderator. Contrasts were tested for equality with the package 'multcomp' (Hothorn, Bretz & Westfall, 2008) in R. For post-hoc tests for between- and within-family contrasts, we ran a meta-analysis on the combined between- and within-family results and included a dummy moderator variable for the between and within-family contrasts, i.e., 0= within-family, 1= between-family. The *V*-matrices are in this case given by adding the Vwithin matrix to the diagonal of the Vbetween matrix, setting all between-within covariances to zero:

$$\begin{pmatrix} V_{between} & 0 \\ 0 & V_{within} \end{pmatrix} \qquad \text{Equation 3.}$$

Transmitted/ Non-transmitted alleles

Just as with the between- and within-family analyses, we first modeled the effects between the transmitted (T-PGS) and the non-transmitted PGSs (NT-PGS) and aggression for each of the 24 data collection waves in a mixed effects model with the package lme4 (Bates, 2015) in R. The mixed effects regression model can be written as:

$$AGGij = \alpha_{0j} + \beta_T * \text{T-PGS} + \beta_{NT} * \text{NT-PGS} + \beta_{3\text{-}v} * x_{3\text{-}v} + \varepsilon_{ij} \qquad \text{Equation 4.}$$

In this notation, AGG*ij* is the aggression outcome of individual *i* in family *j*, the intercept α*0j* is a combination of the overall intercept and the family-level deviation of that intercept, β_T and β_{NT} represent the expected change in the aggression outcome *Y* given a one SD change in the T-PGS and NT-PGS respectively. $\beta_{3\text{-}v}$ represent the expected change in the aggression outcome given a one SD change in the included covariates $x_{3\text{-}v}$ (age, sex, dummy variables for genotyping arrays, and five ancestry-based principal components). The random intercept accounts for any dependence between measures in twins or siblings.

The results from the analyses in the 24 data collection waves were meta-analyzed in a fixed-effects model in Metafor (Viechtbauer, 2010) in R. The meta-analytic method is identical to the between- and within-family PGS meta-analysis described above, with dependence between subsamples accounted for by including the variance-covariance matrix of the sampling errors in the model. See Supplements Table S6.3 and S6.4 for the sample overlap matrix and cross-sample correlation matrix. The same post-hoc tests of between-rater and between- and within-family contrasts were performed for the transmitted and non-transmitted results.

RESULTS

Between- and within-family PGS prediction

After running within- and between-family PGS analyses for each of the 24 data collection waves, we meta-analyzed the results, both per rater and for all measures combined (see Figure 6.1, Table 6.4, and Supplemental Figures S6.1 to S6.6). Results indicate significant associations between aggression in the between- and in the within-family analyses of EL-AGG, EA, and ADHD. The point estimates for the within-family effects were smaller compared to the between-family effects. However, these differences were not significant, indicating no significant confounding of between-family PGS effects by correlated family factors. Rater specific meta-analyses indicate significant between rater differences, most notably between parent report on one hand, and teacher and self-report on the other. The direction of the effects is consistently in the same direction, and differences between raters are small, indicating that sample size is a potential driver of the between rater differences (sample size is largest for parent report, see Table 6.2). For the analyses based on the ADHD PGSs, there were no significant rater differences in the between-family results, but there were significant rater differences in the within-family results. However, the differences are small, and could be solely induced by sample-size/power differences. See Table 6.4 for a complete overview of all standardized regression estimated and confidence intervals.

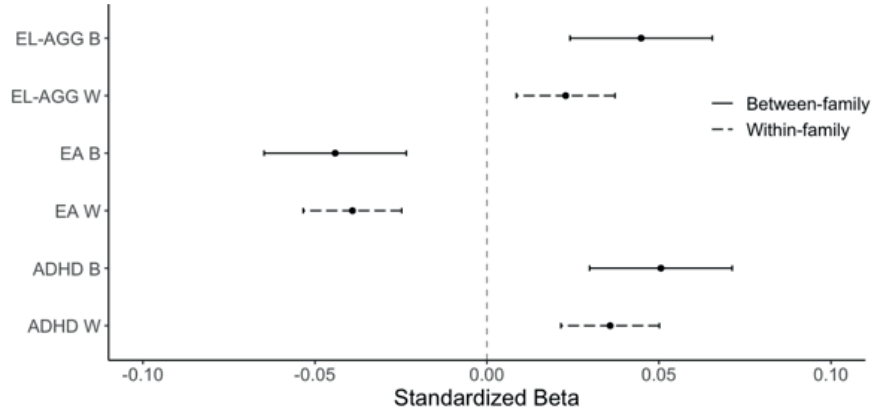


Figure 6.1. Between- and within-family PGS effects with bars depicting 95% confidence intervals

Note: EL-AGG= early-life aggression PGS, EA= educational attainment PGS, ADHD= Attention Deficit Hyperactivity Disorder PGS, B= between-family model, W= within-family model.

Transmitted and non-transmitted PGS prediction

After running the transmitted and non-transmitted PGS analyses for each of the 24 data collection waves, we meta-analyzed the results, both per rater and for all measures combined (see Figure 6.2, Table 6.5, and Supplemental Figures S6.7 to S6.12). Results indicate significant associations between aggression and the transmitted EL-AGG, EA, and ADHD PGSs. The non-transmitted EL-AGG, EA and ADHD PGSs were not significantly associated with the aggression outcome. This indicates no significant contribution of indirect genetic effects from parental rearing environment to offspring aggression. Rater specific meta-analyses indicate significant between rater differences in the transmitted PGSs, most notably between parent report on one hand, and teacher and self-report on the other. Just as in the within-/between-family results, the direction of the rater specific effects is consistently in the same direction, and differences between raters are small, indicating that sample size is a potential driver of the between rater differences (sample size is largest for parent report, see Table 6.3). See Table 6.5 for a complete overview of all standardized regression estimated and confidence intervals.

Table 6.4. Between- and within-family PGS meta-analysis results

EL-AGG						
	Between			Within		
	β	*CI min*	*CI max*	*β*	*CI min*	*CI max*
	0.04***	0.02	0.07	0.02**	0.01	0.04
Rater contrasts						
SR-PR	-0.02	-0.06	-0.02	-0.01	-0.04	0.02
TR-PR	-0.01	-0.06	0.03	-0.01	-0.04	0.02
TR-SR	0.01	-0.04	0.06	0.00	-0.04	0.03
Within-between contrast	-0.02	-0.05	0.00			
EA						
	Between			Within		
	β	*CI min*	*CI max*	*β*	*CI min*	*CI max*
	-0.04***	-0.06	-0.02	-0.04***	-0.05	-0.02
Rater contrasts						
SR-PR	0.07***	0.03	0.11	0.04**	0.01	0.06
TR-PR	0.00	-0.04	0.05	0.02	-0.01	0.05
TR-SR	-0.07**	-0.12	-0.02	-0.02	-0.06	0.01
Within-between contrast	0.01	-0.02	0.03			
ADHD						
	Between			Within		
	β	*CI min*	*CI max*	*β*	*CI min*	*CI max*
	-0.04***	-0.06	-0.02	0.04***	0.02	0.05
Rater contrasts						
SR-PR	-0.01	-0.05	0.03	-0.03*	-0.06	-0.01
TR-PR	0.02	-0.02	0.07	0.00	-0.03	0.03
TR-SR	0.03	-0.01	0.08	0.04*	0.00	0.07
Within-between contrast	-0.01	-0.04	0.01			

Abbreviations: EL-AGG= early life aggression PGS, EA= educational attainment PGS, ADHD= Attention Deficit Hyperactivity Disorder PGS, SR = Self Report, PR = Parent Report, TR = Teacher Report, CI min= confidence interval lower bound, CI max= confidence interval upper bound, β= standardized regression coefficient.
Note: *= p< 0.05, **= p< 0.01, ***= p< 0.001

Table 6.5. Transmitted and non-transmitted PGS meta-analysis results

EL-AGG						
	Transmitted			*Non-transmitted*		
	β	*CI min*	*CI max*	*β*	*CI min*	*CI max*
	0.04***	0.02	0.06	0.01	-0.01	0.03
Rater contrasts						
SR-PR	-0.03	-0.06	0.01	-0.02	-0.06	0.02
TR-PR	-0.01	-0.05	0.02	0.01	-0.03	0.05
TR-SR	0.01	-0.03	0.06	0.03	-0.01	0.08
T-PGS – NT-PGS contrast	0.04**	0.01	0.07			
EA						
	Transmitted			*Non-transmitted*		
	β	*CI min*	*CI max*	*β*	*CI min*	*CI max*
	-0.07***	-0.09	-0.05	-0.01	-0.03	0.01
Rater contrasts						
SR-PR	0.06***	0.03	0.10	0.00	-0.03	0.04
TR-PR	0.02	-0.02	0.05	0.00	-0.03	0.04
TR-SR	-0.05*	-0.09	0.00	0.00	-0.04	0.05
T-PGS – NT-PGS contrast	-0.06***	-0.08	-0.03			
ADHD						
	Transmitted			*Non-transmitted*		
	β	*CI min*	*CI max*	*β*	*CI min*	*CI max*
	0.08***	0.06	0.10	0.00	-0.02	0.02
Rater contrasts						
SR-PR	-0.02	-0.06	0.01	0.01	-0.02	0.05
TR-PR	0.02	-0.01	0.06	0.01	-0.02	0.05
TR-SR	0.05*	0.00	0.09	0.00	-0.04	0.05
T-PGS – NT-PGS contrast	0.09***	0.06	0.11			

Abbreviations: EL-AGG= early life aggression PGS, EA= educational attainment PGS, ADHD= Attention Deficit Hyperactivity Disorder PGS, SR = Self Report, PR = Parent Report, TR = Teacher Report, CI min= confidence interval lower bound, CI max= confidence interval upper bound, β= standardized regression coefficient.
Note: *= p< 0.05, **= p< 0.01, ***= p< 0.001

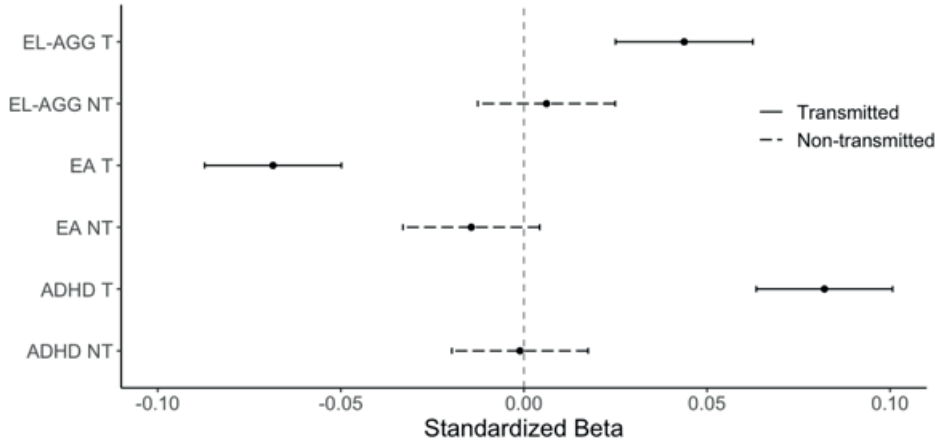


Figure 6.2. Transmitted and non-transmitted PGS effects. with bars depicting 95% confidence intervals

Note: EL-AGG= early-life aggression PGS, EA= educational attainment PGS, ADHD= Attention Deficit Hyperactivity Disorder PGS, T= transmitted PGS model, NT= non-transmitted PGS model.

DISCUSSION

In this study we investigated whether indirect genetic effects play a role in the familial clustering of aggression. Results indicate significant positive direct genetic effects on aggression for the EL-AGG and ADHD PGSs, and a significant negative effect on aggression for the EA PGS. Within-family PGS effect estimates were smaller than the between-family estimates, but these differences were not significant, indicating no significant contribution of indirect genetic effects on aggression. We found no significant contribution of the non-transmitted PGSs (NT-PGS) on aggression, again indicating no significant contribution of indirect genetic effects on aggression. Our study further strengthens the evidence obtained from twin and adoption studies that (direct) genetic effects are the most important driver of familial clustering of aggression.

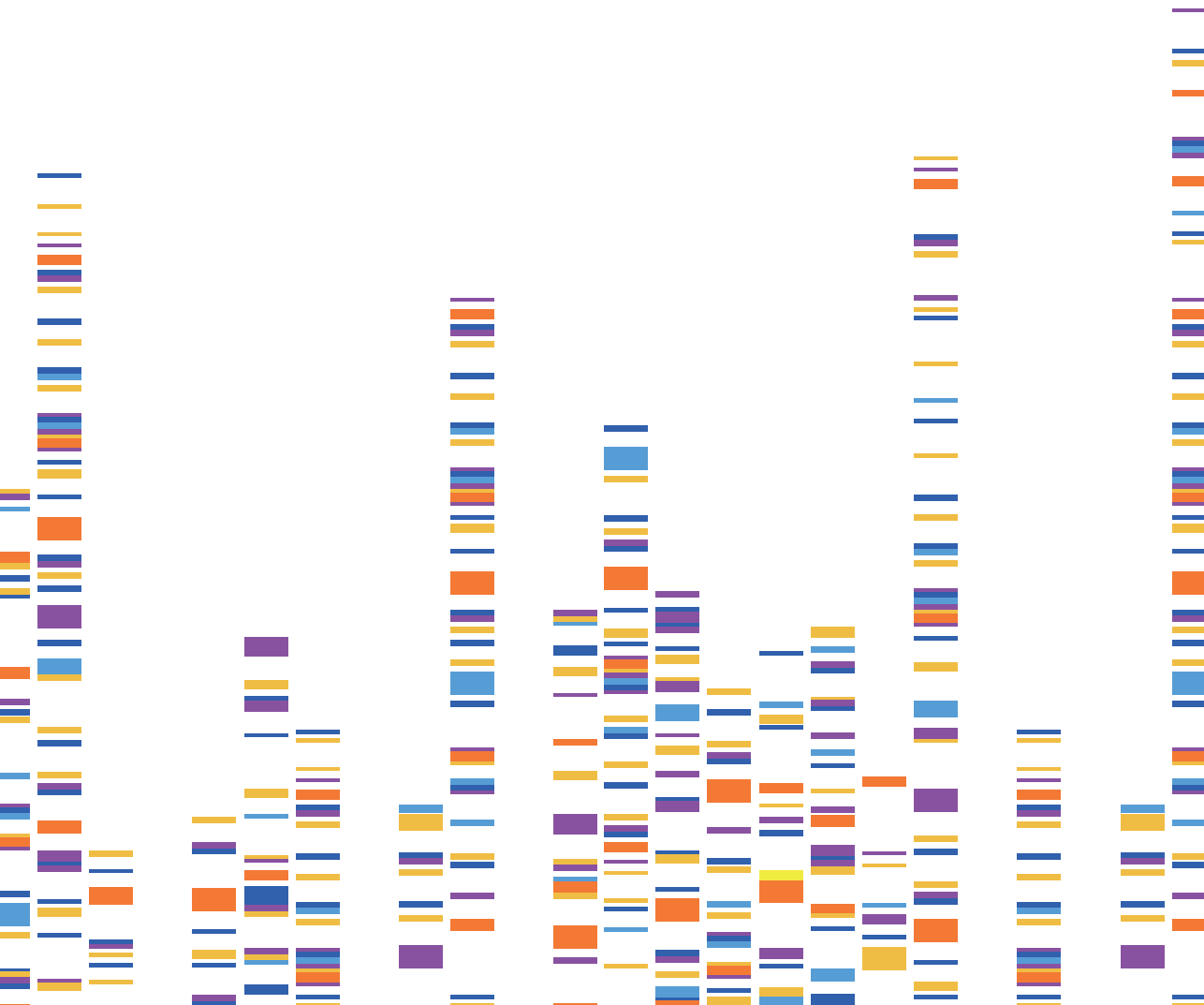
The results may be seen as surprising, because aggression can be hypothesized to be a good example of a trait that may be influenced by indirect genetic effects such as a violent environment (i.e., an environmental manifestation of parents' or siblings' polygenic propensity for aggression). Our results suggest that these family environmental factors are not confounding the genotype-aggression associations. We found direct genetic

effects on aggression, indicating genetic transmission of aggression from parent to offspring. The direct genetic effects indicate a shared genetic basis between EL-AGG, EA, ADHD and aggression, a causal association between these phenotypes, or a genetic effect on another trait that influences all three traits.

We found some slight interrater differences in our results, although all effects were in the same direction across raters. There were clear sample sizes differences between raters, which could have induced (part of) the interrater differences. Another clear difference is the age of the respondents across raters, with teacher and parent ratings covering children only, while self-report also covers adults. This means that it is hard to interpret the interrater differences as being induced by rater-bias, i.e., scoring individuals differently because they see them in different contexts or because they compare them to their siblings or even themselves (Bartels et al., 2003).

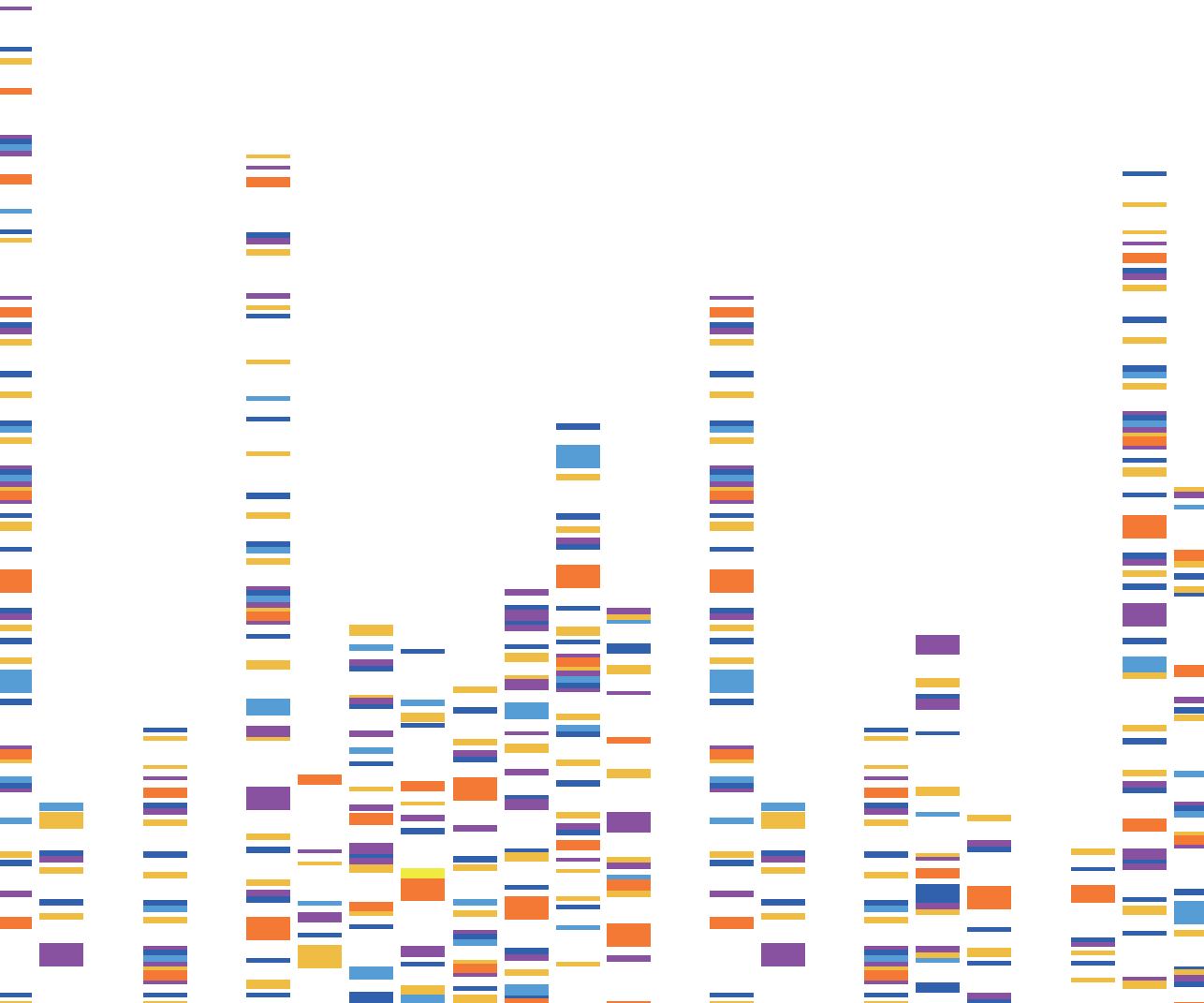
Our study is limited by the statistical power of discovery GWASs. So how sure are we that within- and between family models and transmitted/non-transmitted allele models are able to pick up indirect genetic effects? When looking at the results from the within-/between-family analyses, there is a possibility that we are failing to detect significant differences between the within- and between-family PGS estimates due to a lack of power. The estimates for the within-family effects are slightly smaller than the estimates of the between-family effects. However, the difference between the within- and between-family estimates is smallest for EA, which is the highest powered PGS used in this study. Furthermore, the estimates for the non-transmitted PGS effects were all very close to zero. This leads us to conclude that PGS influences on aggression are mainly direct, and that the contribution of indirect genetic effects seems to be small compared to other behavioral traits such as educational attainment, where significant effects of non-transmitted PGSs were often picked up (Kong et al., 2018; Demange et al., 2020), though not always (De Zeeuw et al., 2020).

In sum, we found no significant contributions of indirect genetic effects on aggression, across different designs. Results indicate significant direct genetic effects on aggression for three PGSs: early-life aggression, educational attainment and attention-deficit hyperactivity disorder. This study provides further evidence that the main drivers of familial clustering and intergenerational transmission of aggression are direct genetic influences that are not confounded by family environment.



7

Summary and Discussion



SUMMARY

The goal of the research described in this dissertation was to investigate the importance of family factors in explaining individual differences in levels of aggression, in explaining whether individuals conform or desist from the overall downward trend we see in aggression over time and across the life-course, and in explaining the continuity of relative aggression levels across the life-course and across generations. We found that the main drivers of familial clustering of aggression and aggression trends are genetic factors. The environment was not found to lead to clustering in families, in contrast, environmental influences mainly lead to differences between family members. Below I will summarize the findings from the studies described in this thesis, and conclude with a general discussion with directions for future research.

A need for interdisciplinary approaches

In Chapter 2, I described the interdependence of potential risk factors, and argue for the need of interdisciplinary approaches and genetically sensitive study designs that help overcome challenges of interdependence. Interdependence of (family) risk factors in this context means that they are hard to disentangle, because they are nested within individuals and contexts (Bronfenbrenner, 1979). When we want to investigate the association between a risk factor and a trait, we are at risk of introducing bias, because this risk factor is likely correlated with a number of other potentially causal factors.

Several study designs are discussed in Chapter 2 that can be used to test (causal) relationships between potential risk factors and aggression in the absence of a varying degree of genetic and/or environmental confounding effects. These include discordant twin or sibling designs and Direction of Causation models (Heath et al., 1993; Duffy & Martin, 1994) as examples of twin and within-family designs, and include Mendelian Randomization (Smith & Ebrahim, 2003; Boutwell & Adams, 2020), within-/between-family PGS designs, and transmitted/ non-transmitted PGS designs (Bates et al., 2018; Kong et al., 2018) as examples of molecular genetic approaches.

These designs are not only of value to researchers with an interest in genetics. On the contrary, especially within-family designs, Direction of Causation models, and Mendelian Randomization designs are eminently suited to study associations between an environmental predictor and

outcome while controlling for varying proportions of genetic confounding and confounding by other unmeasured family factors, especially when the golden standard of causality research – the randomized controlled trial – is not feasible.

We used a within- and between-family design in Chapter 3 to assess clustering of aggression levels and of changes over time and life-course in aggression within families. In Chapter 6 we used a within- and between-family PGS design to investigate contributions of indirect genetic effects, and a transmitted/non-transmitted PGS design (Bates et al., 2018; Kong et al., 2018) to investigate specific indirect genetic effects that are the result of passive gene-environment correlation (prGE).

Aggression is dynamic, and clusters within families

In Chapter 3, the objective was to examine trends in self-reported aggression from 1991 to 2015, to investigate whether these trends apply equally to all individuals, and to explore the extent to which differences in trends over time cluster within families. Forms and mean levels of aggression often vary across the life-course (Tremblay et al, 2004; Tremblay, 2010). Additionally, the prevalence of physical aggression (Frøyland & von Soest, 2018; Pickett et al., 2013) and certain extreme forms of aggression (e.g. murder and violent crime; FBI, 2019) are known to change over time, as society as a whole changes. Using data from 40,400 individuals from 15,437 families that were registered with the Netherlands Twin Register, we found a steady decline in self-reported aggression over time, between 1991 and 2015. We also found that the level of aggression tends to decrease over the life course, as people age. In a novel mixed effects approach in aggression research, we tested whether these trends over time and age applied to all individuals, and whether conforming or desisting from the trends clustered within families. We found that not all individuals follow the downward trend over time or over age to the same extent. There was clear evidence that overall aggression levels, and trends over time and age, cluster within families. In other words, family factors are not only important in explaining individual variation in aggression levels, but also in understanding differences between individuals in time trends. Importantly, the random family slope of the aggression trend (i.e., the familial clustering of conforming or desisting from the overall downward trend we observed) was not correlated with the random family intercept (i.e., the familial clustering of aggression levels), implying that the family factors that cause individual differences in aggression levels are not necessarily the same factors that cause individuals to follow, or desist from

the overall downward trend we observed. Risk factors for aggression that reside in the family system, may not be static but dynamic, changing over time as society as a whole changes.

Genetic continuity of aggression

The goal of Chapter 4 was to test to what extent genetic influences play a role in the stability of relative aggression levels over the life-course. Although the form and level of aggression changes across the life-course, we still see that individuals largely retain their relative positions in aggression levels (Pulkkinen & Pitkänen, 1993; Tuvblad & Baker, 2011). This is interesting because it suggests that individuals' relative propensity for aggression remains stable across situations and life-phases. We assessed whether genetic influences play a role in this stability, by testing to what extent genetic influences that explain individual differences in early-life aggression also explain individual differences across the life-course. In two cohorts from The Netherlands (*N*= 13,471) and Australia (*N*= 5,628), Polygenic Scores (PGSs) were computed based on a genome-wide association study (GWAS) of early-life aggression (Ip et al., 2021). In a novel analytic approach, we ran a mixed effects model for each age (Netherlands: 12-70 years, Australia: 16-73 years), with observations at the focus age weighted as 1, and decaying weights for ages further away. We called this approach a 'rolling weights' model. In The Netherlands, the estimated effect of the PGS was significant and relatively similar from age 12 to age 41, and decreased from age 41 to 70. In Australia, there was a peak in the effect of the PGS around age 40 years. These results are a first indication from a molecular genetics perspective that genetic influences on aggressive behavior that are expressed in childhood continue to play a role up to around age 40. This adds to a growing body of literature that shows that a large part of the stability of aggression across the life course is likely driven by genetic influences (van Beijsterveldt et al., 2003; Gustavson et al., 2020).

Intergenerational transmission of aggression

In Chapter 5 and Chapter 6, the goal was to better understand why family members are alike in their propensity for aggression, as was demonstrated in Chapter 3 and many previous studies (Frisell, Lichtenstein, & Långström, 2011; van de Weijer et al., 2014; Veroude et al., 2016). Broadly speaking, there are two mechanisms through which behavior can be transmitted or continued from parents to offspring: genetic transmission and cultural transmission. Genetic transmission refers to the transmission of behavior through genes that are inherited from parents. Cultural transmission is the transmission

of behavior through the passing on of norms, values, and knowledge from parents to offspring (Cavalli-Sforza & Feldman, 1981). In Chapter 5, based on the differences in genetic resemblance between monozygotic and dizygotic twin pairs, we estimated to what extent differences in rule-breaking behavior, which is closely related to aggression (Bartels et al., 2003), could be explained by genetic and environmental factors. Twin studies can tell us whether the intergenerational transmission of aggression is likely driven primarily by genetic transmission, i.e., heritability inherently implies genetic transmission of behavior, or by cultural transmission, i.e., by shared environmental influences that reflect the maximum effect of the transmission of cultural norms and values from parents to offspring (Keller et.al., 2009). The results indicated that 60 percent of the variance in rule-breaking behavior could be attributed to genetic influences. There was no evidence for shared environmental influences. The remaining 40 percent of the variance was explained by non-shared environmental factors, which lead to differences between family members. These results suggest that intergenerational transmission of rule-breaking behavior is predominantly driven by genetic transmission. Previous studies on aggression show similar results, with heritability estimates varying around 50 percent, and with relatively small contributions of shared environment (Veroude et al., 2016; Odinstova et al., 2019).

In most twin and family studies, we assume that genotype and environment are uncorrelated. If genes (G) and shared environment (C) are correlated, this would lead to an overestimation of C and underestimation of G, if genes (G) and unique environment (E) are correlated, this would lead to an overestimation of G and underestimation of E. This means that the lack of C we find in twin and family studies suggest that genes and shared environment are uncorrelated. However, there are multiple theoretically plausible mechanisms that can induce gene-shared environment correlations. For example assortative mating, population stratification, and passive genotype-environment correlation through environmental manifestations of parents' and siblings' genotypes can all lead to gene-shared environment correlations. Passive genotype-environment correlation may be induced because the environment in which parents rear their children, is partly dependent on their genotypes. This means that children inherit both their genotype and the related environment from their parents, possibly inducing a correlation (Plomin, DeFries & Fulker, 1977; Kendler & Eaves, 1986; Plomin, 2014). Some hypotheses on gene-shared environment correlations can be directly tested in PGS designs. In Chapter 6, we tested for the

presence of indirect genetic effects as a result of gene-shared environment correlation in a within- and between-family PGS design, and directly tested for prGE effects in a transmitted/non-transmitted PGS design (Bates et al., 2018; Kong et al., 2018). We used three PGSs computed from the summary statistics of three large GWASs: early-life aggression (EL-AGG; Ip et al., 2021), Attention-Deficit Hyperactivity Disorder (ADHD; Demontis et al., 2019), and educational attainment (EA; Lee et al., 2018). Results indicated no significant contributions of indirect genetic effects, i.e., no environmental confounding of genetic effects, but did indicate significant direct genetic effects on aggression for all three PGSs. As such, this study provides further evidence that the main drivers of familial clustering and intergenerational transmission of aggression are direct genetic influences.

This study was limited by the statistical power of the discovery GWASs. So how sure are we that within- and between family models and transmitted/non-transmitted PGS models are able to pick up indirect genetic effects? Effect sizes for the within-family analyses were smaller than the between-family estimates, suggesting we might have missed significant indirect genetic effects due to power. However, the difference between the within- and between-family estimates was smallest for EA, which was the highest powered PGS used in this study. Furthermore, the estimates for the non-transmitted PGS effects were all very close to zero. This leads us to conclude that genetic influences on aggression are mainly direct, and that the contribution of indirect genetic effects seems to be small compared to other behavioral traits such as educational attainment, where significant effects of non-transmitted PGSs were often picked up (Kong et al., 2018; Demange et al., 2020), though not always (de Zeeuw et al., 2020).

GENERAL DISCUSSION

The studies bundled in this dissertation tell a clear story: genetic factors are the main drivers of familial resemblance in aggression, and environmental influences predominantly lead do differences in the propensity for aggression. So what are the mechanisms that drive these findings?

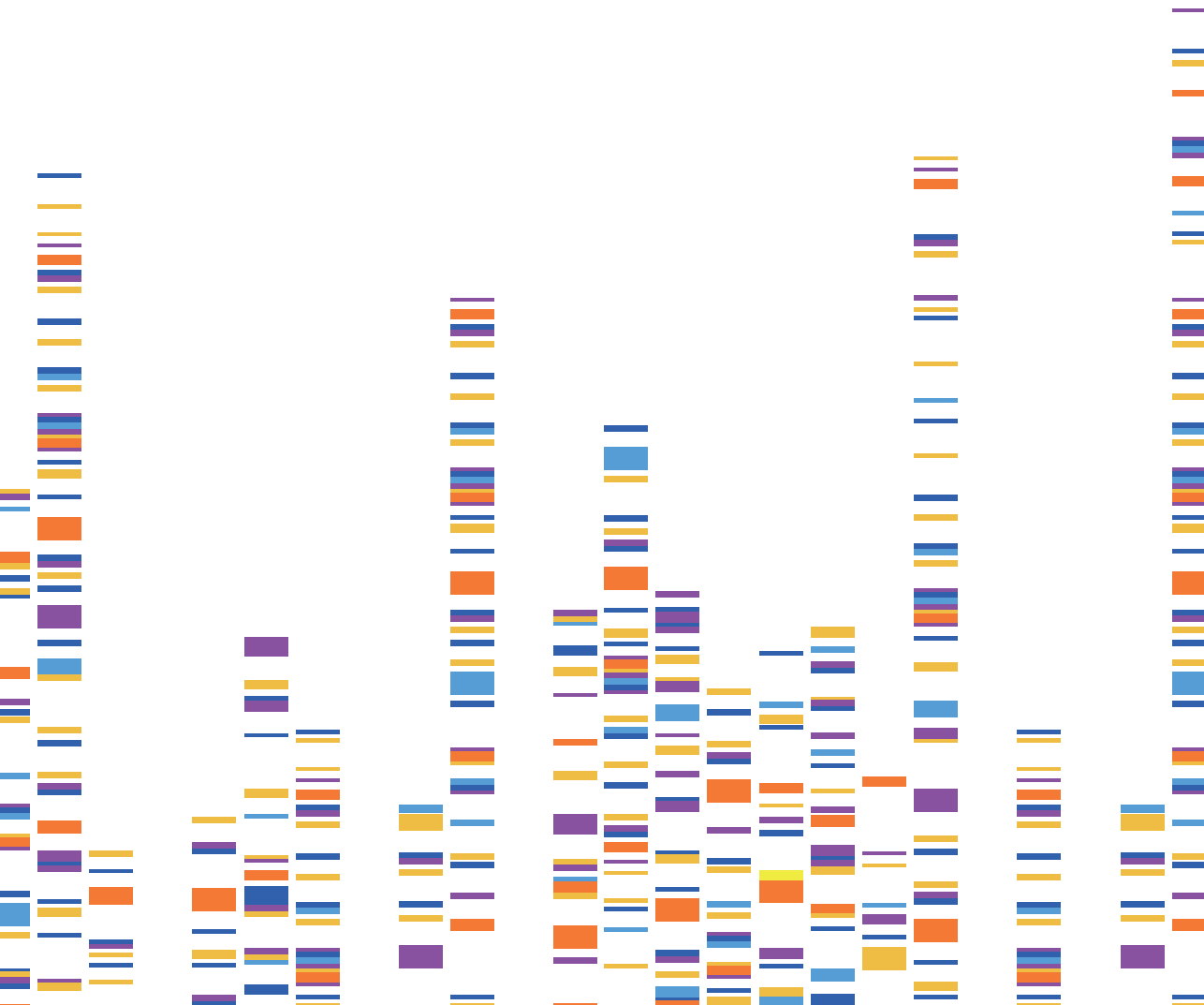
Much about the functional pathways behind genetic influences on aggression remains unclear. Aggression is highly polygenic, meaning that a multitude of genetic variants influence aggression, each with a very small effect size. This suggests that the total genetic effects run through many different mechanisms. The small effect sizes make it difficult to detect influences of single functional genetic variants that could provide clues on causal biological pathways. However, there are various theories on functional mechanisms that could explain genetic influences on aggression. Raine (2008) argues that an important functional pathway is through genetic influences on brain structures. For example the MAOA gene, which has often been mentioned in relation to aggression (see for an overview Odintsova et al., 2019), codes for an enzyme that breaks down serotonin, and is associated with the volume of the amygdala, anterior cingulate, and orbitofrontal cortex (Meyer-Lindenberg et al., 2006). These brain structures are associated with emotional, cognitive, and behavioral deficits, revealing a potential causal pathway from genes to aggression (Raine, 2008). However, not a single genetic variant-aggression association has replicated in GWAS (Odintsova et al., 2019), illustrating the difficulty of identifying and proving causal gene-aggression pathways. Another potential pathway from genes to aggression could run through biomarkers. For example lipid levels (Hagenbeek et al., 2018), resting heart rate (Raine et al., 2014; Latvala et al., 2016), and testosterone levels (Brain & Susman, 1997; Ramirez, 2003) have all been suggested as potential causal influences on aggression. However, Ip et al. (2021) found very little evidence to suggest that the pathway from genes to aggression runs through these biomarkers: genetic correlations between early-life aggression and each of these biomarkers were small. Whatever the causal mechanisms may be, we will need high powered GWASs to prove hypotheses on causal mechanisms, or to find clues that may lead to new hypotheses. An important requirement for an increase in the statistical power of GWASs is more and wider available data. One exciting project that is currently underway in The Netherlands is the ODISSEI (Open Data Infrastructure for Social Science and Economic Innovations) project (odissei-data.nl). ODISSEI is a project funded by a NWO roadmap grant,

and, among other tasks, aims to enhance the infrastructure that facilitates sharing and linking of data among researchers and for example Statistics Netherlands, the population registry of The Netherlands. De Zeeuw and colleagues (2021) showed that sensitive cohort and population registry data can be safely and successfully linked on the ODISSEI Secure Supercomputer platform. This will allow researchers to run GWASs on traits that may not have been directly collected in their cohort, potentially increasing sample sizes and broadening the scope of traits to study.

In our studies, we found little evidence for shared environmental influences on aggression, i.e., influences that are not genetic and cause family members to be more similar in their propensity for aggression compared to two random unrelated individuals. We demonstrated that the genetic effects were not significantly confounded by effects of a potentially correlated family environment. This begs the question whether associations between environmental risk factors and aggression may be confounded by genetic effects. This hypothesis can often be directly tested, as described in Chapter 2. A design that is attainable for researchers across research fields, and that is well suited to control for a varying range of (genetic) confounding, is a discordant (MZ) twin and/or sibling design. In order to be able to apply such designs, it is necessary that researchers more often collect data on complete families rather than unrelated individuals. One of the tasks that ODISSEI will work on in the near future is to unlock the full potential of within-family analyses in population registry data in The Netherlands, by defining a thorough network of family ties, including the correct classifications of genetic resemblance, and make this available to all researchers in The Netherlands. I believe this is an important next step in aggression research, because it allows for more direct testing of hypotheses on causal associations. A better understanding of causal mechanisms behind the development of aggression and related traits may benefit the efficacy of prevention and intervention strategies. That being said, it is important to note that if an association is confounded by genetic effects, this does not necessarily mean that intervening on that factor is useless. As with all research, the results are dependent on the context in which they were studied. By intervening on certain family factors, this context may change, which may have beneficial effects on aggression.

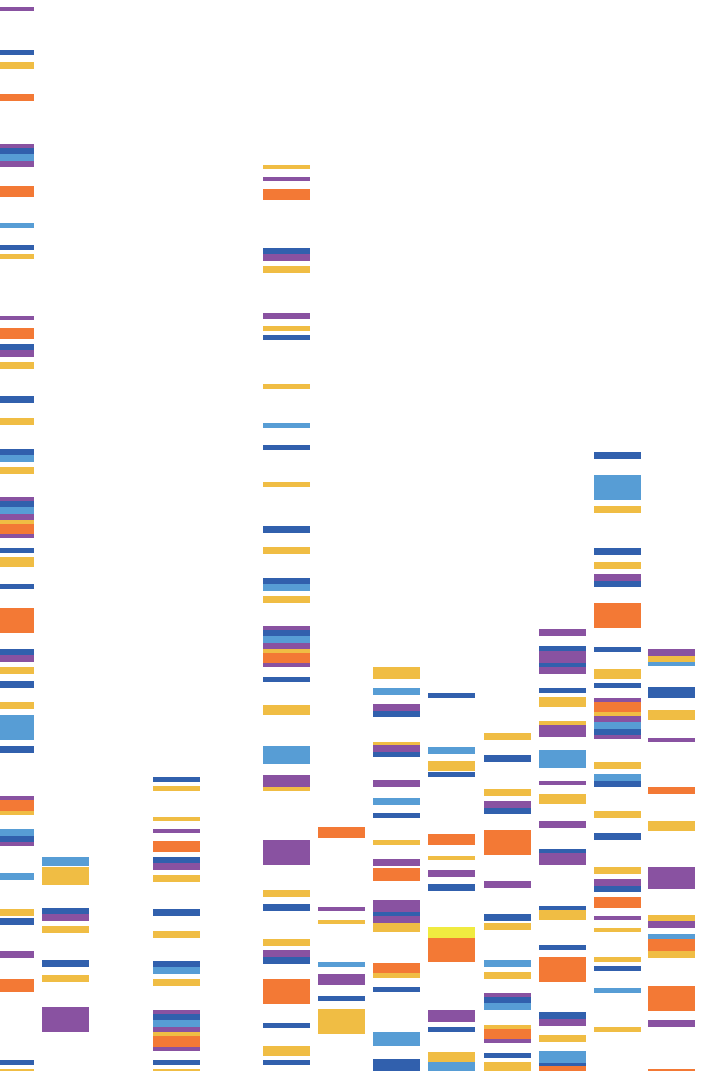
Much of this dissertation has focused on shared environmental influences, i.e., to what extent do environmental influences lead to similarities within families. While we have found little to no evidence for these shared

environmental influences, this does not mean that the environment has no impact on aggression. On the contrary, we found that the environment explained 40% of the variance in rule-breaking behavior, while previous research has found on average 50% of the variance in aggression to be attributable to environmental influences (Veroude et al., 2016; Odintsova et al., 2019) These environmental influences caused differences in outcomes within families, to such an extent that environmental influences led to as many differences between family members as they would between any two random individuals from the population. So what are these 'unique' environmental influences that explain a large part of the individual differences in aggression? First of all, the unique environmental influences in twin and family models inherently include all measurement error. The remaining environmental influences include the main or direct effects of personal experiences on aggression, and all interaction effects with genotype and shared environmental factors. Asbury, Dunn, and Plomin (2006) took to a qualitative approach to assess what these unique environmental influences or personal experiences may be, in their case for childhood anxiety. From 19 interviews they identified negative school experiences, twin comparison, illness and accidents, traumatic neonatal life events, parent–child relationships, and peer rejection as potential unique environmental influences. I expect that their findings may also apply to aggression, in the sense that the personal experiences that influence aggression likely vary between individuals. Moreover, just as with genes, I expect that the effect of most personal experiences is small. As such, well powered within-family designs, as described in Chapter 2, are needed to be able to detect such unique environmental influences on a population level.



8

Nederlandstalige Samenvatting en Discussie



SAMENVATTING

Het doel van het onderzoek dat staat beschreven in deze dissertatie was om inzicht te krijgen in de rol van familiefactoren bij het verklaren van individuele verschillen in agressie, bij het volgen of afwijken van de geobserveerde agressie trends over tijd en over de levensloop, en bij het verklaren van de stabiliteit van agressie over de levensloop en tussen generaties. De resultaten lieten zien dat genetische factoren de grootste drijfveren achter familiaire clustering van agressie en agressie trends waren. Omgevingsfactoren leidden niet tot clustering van agressie binnen families, maar verklaren wel voor een groot deel verschillen tussen familieleden. Hieronder vat ik de Hoofdstukken in deze dissertatie samen, en sluit ik af met een algemene discussie.

De noodzaak van een interdisciplinaire benadering

In Hoofdstuk 2 werd gepleit voor een interdisciplinaire aanpak met genetisch geïnformeerde onderzoeksdesigns die helpen de moeilijkheden het hoofd te bieden die gepaard gaan met de onderlinge afhankelijkheid van risico factoren. Risico factoren zijn genest in individuen en omgeving, waardoor ze lastig te ontrafelen zijn (Bronfenbrenner, 1979). Wanneer we een associatie tussen een risico factor en een eigenschap willen testen, lopen we het risico dat er bias optreedt omdat de risico factor mogelijk correleert met andere potentieel causale factoren.

In Hoofdstuk 2 werden verschillende onderzoeksdesigns besproken, met een specifieke focus op het testen van (causale) associaties in (gedeeltelijke) afwezigheid van confounding effecten. Relevante voorbeelden van tweeling en familie designs zijn discordante tweeling of broers/zussen designs en Direction of Causation designs (Heath et al., 1993; Duffy & Martin, 1994), voorbeelden van moleculaire genetica benaderingen zijn Mendelian Randomization (Smith & Ebrahim, 2003; Boutwell & Adams, 2020), binnen-familie polygenetische score (PGS) designs, en overgedragen/niet-overgedragen PGS designs (Bates et al., 2018; Kong et al., 2018).

Deze designs zijn niet alleen interessant voor onderzoekers met een specifieke interesse in genetische invloeden. Integendeel, vooral binnen-familie designs, Direction of Causation modellen, en Mendelian Randomization zijn bij uitstek geschikt om associaties tussen een omgevingsfactor en uitkomstmaat te onderzoeken terwijl je controleert

voor een variërende mate van genetische confounding en confounding door andere ongemeten familiefactoren, zeker wanneer de gouden standaard voor causaliteit onderzoek – de randomized controlled trial – niet uitvoerbaar is.

In Hoofdstuk 3 maakten we gebruik van een binnen- en tussen-familie design om clustering van agressie en clustering van trends in agressie binnen families te onderzoeken. In Hoofdstuk 6 gebruikten we een binnen- en tussen-familie design om de invloed van indirecte genetische effecten te onderzoeken, en een overgedragen/niet-overgedragen PGS design (Bates et al., 2018; Kong et al., 2018) om indirecte genetische effecten die het resultaat zijn van passieve gen-omgeving correlatie (prGE) te onderzoeken.

Agressie is dynamisch, en clustert binnen families

Het doel van Hoofdstuk 3 was om trends in zelf-gerapporteerde agressie van 1991 tot 2015 te onderzoeken, om te kijken of deze trends in gelijke mate gelden voor alle individuen, en om te kijken of verschillen in trends clusteren binnen families. Over het algemeen veranderen de mate en vorm van agressie over de levensloop (Tremblay et al, 2004; Tremblay, 2010). Bovendien verandert ook de prevalentie van fysieke agressie (Frøyland & von Soest, 2018; Pickett et al., 2013) en bepaalde extreme vormen van agressie (e.g. moord en geweldsdelicten; FBI, 2019) over tijd, als gevolg van veranderingen in de samenleving. Door gebruik te maken van data van 40,400 individuen uit 15,437 families geregistreerd in het Nederlands Tweelingen Register, vonden we een duidelijk dalende trend in agressie tussen 1991 en 2015, en over de levensloop. Met een voor agressie onderzoek nieuw 'mixed effects' design testten we of alle personen in de steekproef de geobserveerde neerwaartse trend over de levensloop en over tijd volgden, en of het volgen of afwijken van de trend clustert binnen families. De resultaten lieten duidelijk zien dat zowel de prevalentie van agressie, als de trends over leeftijd en tijd, clusterden binnen families. In andere woorden, familie factoren zijn niet alleen belangrijk bij het begrijpen van individuele verschillen in agressie, maar ook bij het begrijpen waarom de een wel, en de ander niet een afname in agressie laat zien. Ook van belang is dat de random slope voor agressie over tijd (i.e., de familiaire clustering van het volgen of afwijken van de algemene trend over tijd) niet correleerde met het random intercept voor agressie (i.e., de familiaire clustering van agressie). Dit impliceert dat de familiefactoren die zorgen voor de clustering van agressie binnen families, niet per definitie dezelfde factoren zijn die een rol spelen bij de toename of afname van agressie over tijd. De risicofactoren die gedeeld worden binnen families zijn mogelijk dynamisch, en kunnen veranderen wanneer de maatschappij verandert.

Genetische continuïteit van agressie

Het doel van Hoofdstuk 4 was om te testen of genetische invloeden die op jonge leeftijd een rol spelen in agressie, ook een rol spelen bij agressie op latere leeftijd, en daarmee deels de stabiliteit van de predispositie voor agressie verklaren. Hoewel de vorm en prevalentie van agressie verandert over de levensloop, behouden personen gedurende hun leven over het algemeen hun relatieve positie in de mate van agressie (Pulkkinen & Pitkänen, 1993; Tuvblad & Baker, 2011). Dit is interessant omdat het suggereert dat individuen hun relatieve predispositie houden, onafhankelijk van de situatie of levensfase. In twee cohorten uit Nederland (*N*= 13,471) en Australië (*N*= 5,628) maakten we polygenetische scores (PGSen) gebaseerd op een genoom-wijde associatie studie (GWAS) van agressie in kinderen en jongeren (Ip et al., 2021). Met een nieuwe analytische aanpak, modelleerden we de invloed van de PGSen op agressie voor elke leeftijd (Nederland: 12-70 jaar, Australië: 16-73 jaar), met de focus leeftijd gewogen als 1, en afnemende gewichten voor leeftijden verder weg. We noemden dit een 'rolling weights' model. De geschatte invloed van de PGS was in Nederland redelijk gelijk van leeftijd 12 tot 41, en nam af van leeftijd 41 tot 70. In Australië was een piek te zien in de invloed van de PGS rond leeftijd 40. Deze resultaten zijn een eerste indicatie vanuit een moleculair genetisch perspectief dat genetische invloeden op agressie bij kinderen en jongeren, ook nog een rol spelen in latere fases in de levensloop. Dit sluit aan bij eerdere studies die lieten zien dat een belangrijk deel van de stabiliteit van agressie over de levensloop waarschijnlijk gedreven wordt door genetische invloeden (e.g. van Beijsterveldt et al., 2003; Gustavson et al., 2020).

Intergenerationele overdracht van agressie

Het doel van Hoofdstuk 5 en 6 was om beter inzicht te krijgen in waarom familieleden op elkaar lijken in hun predispositie voor agressie, zoals werd aangetoond in Hoofdstuk 3 en in eerder onderzoek (Frisell, Lichtenstein, & Långström, 2011; van de Weijer et al., 2014; Veroude et al., 2016). Grofweg zijn er twee mechanismen waardoor gedrag kan worden overgedragen van ouder op kind: genetische transmissie en culturele transmissie. Genetische transmissie duidt op de overdracht van gedrag door het erven van genen. Culturele transmissie duidt op de overdracht van gedrag door het overnemen van normen, waarden, en kennis (Cavalli-Sforza & Feldman, 1981). In Hoofdstuk 5 schatten we het relatieve belang van genetische en omgevingsinvloeden op norm-overschrijdend gedrag, gebaseerd op de verschillen in genetische overeenkomst tussen een- en twee-eiige tweelingen. Norm-overschrijdend gedrag is nauw verbonden

met agressie (Bartels et al., 2003). Tweeling studies kunnen een indicatie geven of intergenerationele overdracht vooral gedreven wordt door genetische transmissie of door culturele transmissie. Immers, genetische invloeden duiden op genetische transmissie van gedrag, terwijl gedeelde omgevingsinvloeden een schatting geven van het maximale effect van culturele transmissie (Keller et al., 2009). De resultaten lieten zien dat 60 procent van de variantie in norm-overschrijdend gedrag verklaard kan worden door genetische invloeden, ook wel de geschatte erfelijkheid genoemd. Er waren geen aanwijzingen voor gedeelde omgevingsinvloeden. Deze resultaten suggereren dat intergenerationele overdracht vooral plaatsvindt via genetische transmissie. Eerdere studies laten eenzelfde beeld zien met betrekking tot agressie, met erfelijkheid schattingen variërend rond de 50 procent, en relatief kleine invloeden van de gedeelde omgeving (Veroude et al., 2016; Odinstova et al., 2019).

In de meeste tweeling en familie studies nemen we aan dat genotype en omgeving niet correleren. Wanneer genen (G) en gedeelde omgeving (C) correleren, leidt dit tot een overschatting van C en onderschatting van G, wanneer genen (G) en unieke omgeving (E) correleren, leidt dit tot een overschatting van G en onderschatting van E. De afwezigheid van C in tweelingstudies suggereert daarom dat G en C niet correleren. Echter, er zijn een aantal theoretisch plausibele mechanismen die tot een correlatie tussen G en C kunnen leiden, waaronder assortative mating, populatie stratificatie, en passieve genotype-omgeving correlatie (prGE). Passieve genotype-omgeving correlatie kan ontstaan doordat de omgeving waarin ouders hun kinderen opvoeden deels afhankelijk is van hun genotype. Omdat kinderen zowel hun genotype als de omgeving van hun ouders erven, ontstaat er mogelijk een correlatie tussen hun genotype en omgeving (Plomin, DeFries & Fulker, 1977; Kendler & Eaves, 1986; Plomin, 2014). Bepaalde hypotheses over gen-gedeelde omgeving correlaties kunnen getest worden in PGS designs. In Hoofdstuk 6 testten we of er sprake was van indirecte genetische effecten door correlaties tussen de (familie)omgeving en genotype in een binnen-en tussen-familie design, en testten we direct voor de aanwezigheid van prGE in een overgedragen/niet-overgedragen PGS design (Bates et al., 2018; Kong et al., 2018). We gebruikten hiervoor drie verschillende PGSen, gebaseerd op de resultaten van drie grootschalige genoom-wijde associatie studies (GWAS): 'early-life' agressie (EL-AGG; Ip et al., 2021), Attention-Deficit Hyperactivity Disorder (ADHD; Demontis et al., 2019), en opleidingsniveau (EA; Lee et al., 2018). De resultaten lieten geen significante invloed van indirecte genetische effecten zien, oftewel

geen indicatie van confounding door de (familie)omgeving. Wel waren er significante directe genetische invloeden van alle drie de PGSen. Daarmee sterkt deze studie de conclusie uit tweelingstudies dat intergenerationele overdracht van agressie vooral komt door genetische transmissie.

Deze studie is gelimiteerd door de statistische power van de GWASen waarop de PGSen waren gebaseerd. Hoe zeker zijn we daarom dat onze binnen- en tussen-familie modellen indirecte effecten kunnen detecteren? Uit de resultaten bleek dat binnen-familie effectgroottes kleiner waren dan de tussen-familie effectgroottes, wat zou kunnen duiden op indirecte effecten die we in onze studie niet konden detecteren. Echter, de verschillen tussen de binnen-family en tussen-family schattingen waren het kleinst voor EA, welke gebaseerd waren op de GWAS met de meeste power. Daarnaast waren de effecten van de niet-overgedragen PGSen alle nagenoeg nul. Dit leidt tot de conclusie dat genetische invloeden op agressie vooral directe invloeden zijn, en dat de invloed van indirecte genetische effecten kleiner lijkt dan bij andere eigenschappen zoals EA, waarbij wel vaak indirecte effecten werden gevonden (Kong et al., 2018; Demange et al., 2020), maar niet altijd (de Zeeuw et al., 2020).

ALGEMENE DISCUSSIE

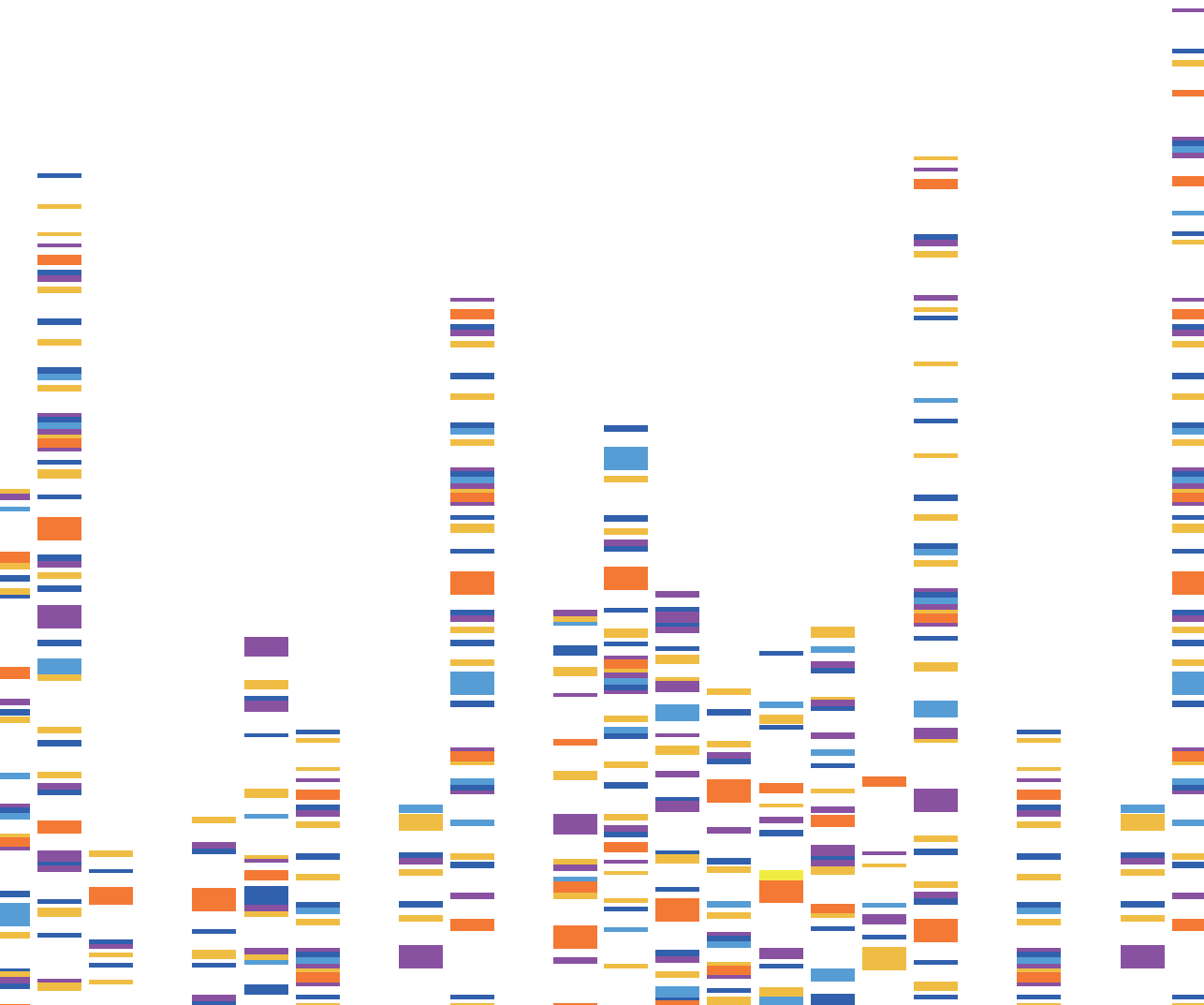
De studies die zijn gebundeld in deze dissertatie laten een duidelijk beeld zien: genetische invloeden zijn de voornaamste drijvende kracht achter familiaire overeenkomsten in agressie, en omgevingsfactoren leiden voornamelijk tot verschillen in de mate van agressie. Maar wat zijn nou de mechanismen achter deze bevinding?

Over de functionele biologische mechanismen achter de invloed van genen op agressie is nog veel onduidelijk. Agressie is zeer polygenetisch, wat betekent dat vele genetische varianten invloed hebben op agressie, elk met een zeer kleine effectgrootte. Dit suggereert dat de totale genetische invloeden via vele verschillende mechanismen lopen. De kleine effectgroottes maken het moeilijk om specifieke functionele genetische varianten te detecteren die mogelijk kunnen helpen bij het identificeren van causale biologische mechanismen. Er zijn echter wel verschillende theorieën over mogelijke manieren waarop genen agressief gedrag kunnen beïnvloeden. Raine (2008) bepleit dat een belangrijk mechanisme loopt via de invloed van genen op verschillende delen van het brein. Bijvoorbeeld, het MAOA gen, vaak genoemd in relatie tot agressie (zie voor een overzicht Odintsova et al., 2019), codeert voor een enzym dat serotonine afbreekt, en is geassocieerd met het volume van de amygdala, de cortex cingularis anterior, en de orbitofrontale cortex (Meyer-Lindenberg et al., 2006). Deze hersengebieden zijn eerder geassocieerd met emotionele, cognitieve, en gedragskenmerken (Raine, 2008). Echter, geen enkele associatie tussen een individuele genetische variant en agressie is tot op heden gerepliceerd in GWASen (Odintsova et al., 2019), wat de moeilijkheid van het identificeren en bewijzen van gen-agressie mechanismen aantoont. Een ander potentieel mechanisme van genen naar agressie kan plaatsvinden via biomarkers. Onder andere lipide niveaus (Hagenbeek et al., 2018), rusthartslag (Raine et al., 2014; Latvala et al., 2016), en testosteron niveaus (Brain & Susman, 1997; Ramirez, 2003) zijn genoemd als mogelijke causale invloeden op agressie. Echter, Ip en collega's (2021) vonden weinig tot geen bewijs dat deze biomarkers gen-agressie associaties kunnen helpen verklaren: genetische correlaties tussen early-life agressie en de hierboven genoemde biomarkers waren klein. Wat de exacte causale mechanismen ook zijn, er zullen GWASen nodig zijn met hoge statistische power om bewijs te kunnen vinden voor hypotheses over causale mechanismen, of om aanwijzingen voor nieuwe hypotheses te kunnen vinden. Een belangrijke voorwaarde voor grotere GWASen is meer en wijder beschikbare data. Een interessant project

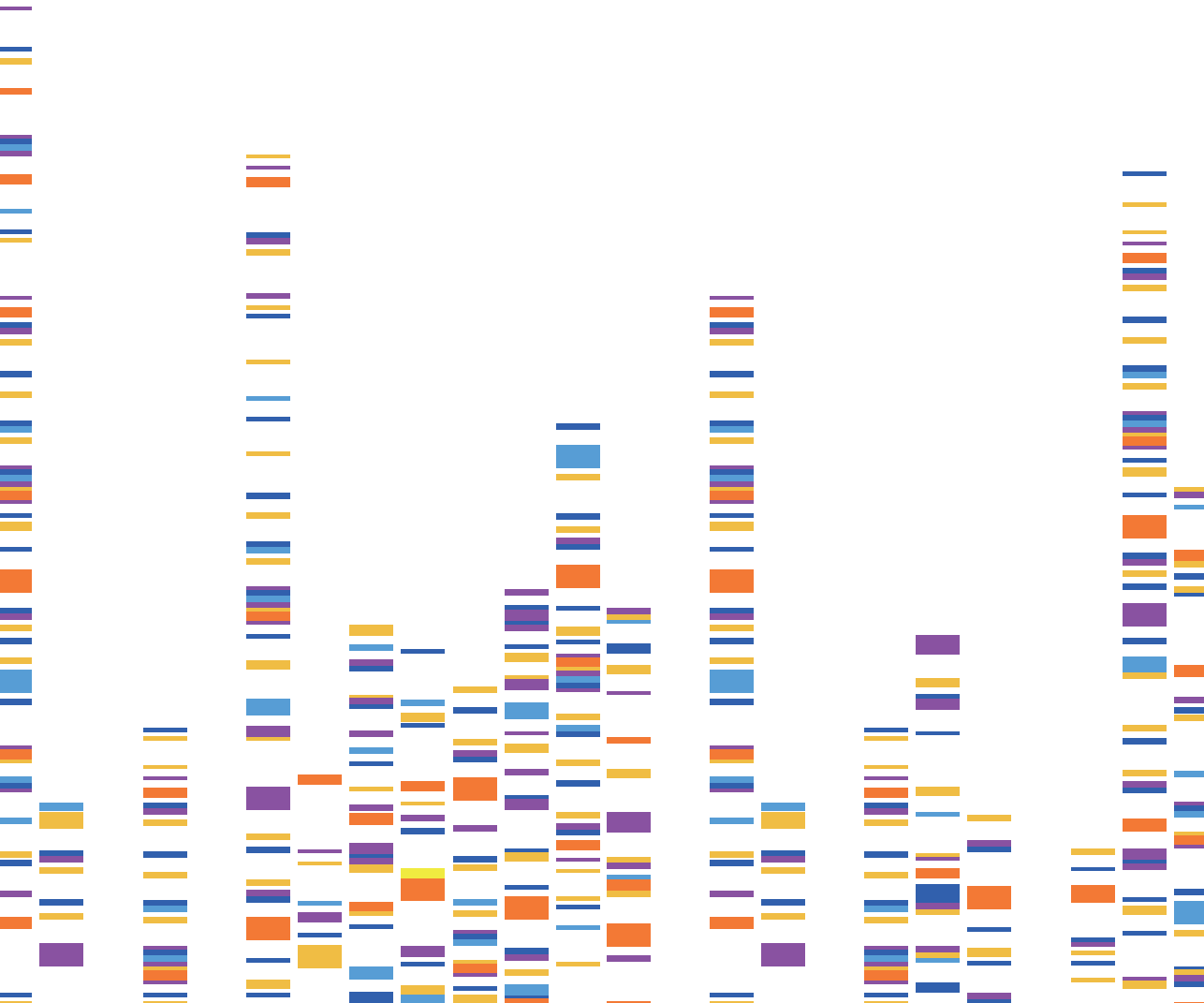
dat momenteel loopt in Nederland is het ODISSEI (Open Data Infrastructure for Social Science and Economic Innovations) project (odissei-data.nl). ODISSEI is een project dat gefinancierd wordt door een NWO roadmap beurs, en heeft onder andere als doel de infrastructuur voor het delen en linken van data tussen onderzoekers en bijvoorbeeld CBS te verbeteren. De Zeeuw en collega's (2021) lieten zien dat het mogelijk is om cohort en CBS data veilig en succesvol te linken op het ODISSEI Secure Supercomputer platform. Dit biedt onderzoekers de mogelijkheid om GWASen te doen met eigenschappen waarvoor zij niet zelf bij hun respondenten data hebben verzameld, wat de steekproeven en variëteit van uitkomsten enorm kan vergroten.

We vonden in ons onderzoek weinig aanwijzingen voor invloeden van de gedeelde omgeving op agressie, i.e., invloeden die niet genetisch zijn en leiden tot meer gelijkheid in agressie uitkomsten tussen familieleden in vergelijking tot niet-familieleden. We lieten bovendien zien dat genetische effecten niet confounded waren door effecten van een mogelijk gecorreleerde familie-omgeving. Dit werpt de vraag op of associaties tussen omgevings-risicofactoren en agressie mogelijk confounded zijn door genetische invloeden. Deze hypothese kan vaak direct getest worden in een genetisch sensitief onderzoeksdesign, zoals beschreven in Hoofdstuk 2. Een design dat relatief makkelijk geïmplementeerd kan worden door onderzoekers uit verschillende vakgebieden, en geschikt is voor het testen van een associatie terwijl je controleert voor een bepaalde mate van confounding door genetische invloeden en ongemeten gedeelde omgevingsinvloeden, is een discordant tweeling of broers en zussen design. Om zulke designs toe te kunnen passen is het van belang dat onderzoekers data verzamelen van families, in plaats van individuen uit verschillende families. Een taak waar ODISSEI zich de komende tijd op richt is om een netwerk van familierelaties in Nederland te definiëren, inclusief de genetische overeenkomsten, en dit beschikbaar te maken voor alle onderzoekers in Nederland. Dit zal het voor onderzoekers makkelijker maken om bepaalde hypotheses over causale verbanden te testen, wat kan leiden tot effectievere preventie en interventie strategieën. Dat gezegd hebbende, het is belangrijk om te realiseren dat wanneer een associatie met een risicofactor confounded is door genetische invloeden, dit niet per definitie betekent dat ingrijpen op die risicofactor kansloos is. Zoals met alle studies, zijn de resultaten afhankelijk van de context waarin ze onderzocht zijn. Door in te grijpen op een bepaalde factor kan de context veranderen, wat invloed kan hebben op de prevalentie van agressie.

Een belangrijk deel van het onderzoek in deze dissertatie was gericht op gedeelde omgevingsinvloeden, i.e., tot in hoeverre leidt een gedeelde omgeving tot gelijke uitkomsten in agressie tussen familieleden. Hoewel we weinig tot geen bewijs vonden voor gedeelde omgevingsinvloeden, betekent dit niet dat de omgevingen helemaal geen invloed heeft op agressie. Integendeel, we vonden dat 40% van de variantie in norm-overschrijdend gedrag toe te schrijven was aan omgevingsinvloeden, en eerder onderzoek liet zien dat gemiddeld 50% van de variantie in agressie toe te schrijven is aan omgevingsinvloeden. Deze omgevingsinvloeden leidden tot verschillen binnen families, in zoverre dat omgevingsinvloeden tot bijna net zoveel verschillen leidden binnen families als tussen twee willekeurige individuen in de populatie. Wat zijn dan precies deze 'unieke' omgevingsinvloeden die een dergelijk grote rol spelen in de individuele verschillen in agressie? Ten eerste behelst de unieke omgeving alle meetfouten. De resterende unieke omgevingsinvloeden zijn de primaire effecten van persoonlijke ervaringen, en alle interactie-effecten met persoonlijke ervaringen, waaronder interacties tussen de persoonlijke ervaringen en genotype, en tussen de persoonlijke ervaringen en de gedeelde omgeving. Asbury, Dunn, en Plomin (2006) onderzochten met een kwalitatieve aanpak wat de specifieke unieke omgevingsfactoren/persoonlijke ervaringen mogelijk kunnen zijn, in dit geval bij angst in kinderen. Op basis van 19 interviews identificeerden ze negatieve school ervaringen, tweeling vergelijkingen, ziekte en ongelukken, traumatische neonatale life-events, ouder–kind relaties, en afwijzing door leeftijdsgenoten als potentiële unieke omgevingsinvloeden. Ik verwacht dat vergelijkbare resultaten zouden worden gevonden bij agressie, in de zin dat persoonlijke ervaringen die invloed hebben op agressie waarschijnlijk variëren tussen individuen. Net als met genetische invloeden, verwacht ik dat effectgroottes van de individuele persoonlijke ervaringen vaak klein zullen zijn. Dit betekent dat grootschalige binnen-familie designs, zoals besproken in Hoofdstuk 2, nodig zijn om deze unieke omgevingsinvloeden te detecteren op populatieniveau.



9
References



REFERENCES

Abdellaoui, A., Hottenga, J.-J., Knijff, P. de, Nivard, M. G., Xiao, X., Scheet, P., Brooks, A., Ehli, E. A., Hu, Y., Davies, G. E., Hudziak, J. J., Sullivan, P. F., van Beijsterveldt, T., Willemsen, G., de Geus, E. J., Penninx, B. W. J. H., & Boomsma, D. I. (2013). Population structure, migration, and diversifying selection in The Netherlands. *European Journal of Human Genetics, 21*(11), 1277–1285. https://doi.org/10.1038/ejhg.2013.48

Achenbach, T. M., Ivanova, M. Y., & Rescorla, L. A. (2017). Empirically based assessment and taxonomy of psychopathology for ages 1½-90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry, 79*, 4-18. https://doi.org/10.1016/j.comppsych.2017.03.006

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

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

Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA Adult Forms & Profiles.* University of Vermont, Research Center for Children, Youth, & Families, Burlington.

Adams, C. D., & Boutwell, B. B. (2020). Adventurousness cuts both ways: A Mendelian randomization of adventurousness on 10 cognitive and behavioral traits. *medRxiv*. https://doi-org.vu-nl.idm.oclc.org/10.1101/2020.03.12.20034918.

Adrian, R. (2008). From genes to brain to antisocial behavior. *Current Directions in Psychological Science, 17*(5), 323-328. https://doi.org/10.1111/j.1467-8721.2008.00599.x

Alink, L. R. A., Mesman, J., Van Zeijl, J., Stolk, M. N., Juffer, F., Koot, H. M., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2006). The early childhood aggression curve: Development of physical aggression in 10- to 50-month-old children. *Child Development. 77*(4), 954-966.

Asbury, K., Dunn, J., & Plomin, R. (2006). The use of discordant MZ twins to generate hypotheses regarding non-shared environmental influence on anxiety in middle childhood. *Social Development, 15*(3), 564-570.

Asimit, J. L., Hatzikotoulas, K., McCarthy, M., Morris, A. P., & Zeggini, E. (2016). Trans-ethnic study design approaches for fine-mapping. *European Journal of Human Genetics, 24*(9), 1330-1336. https://doi.org/10.1038/ejhg.2016.1

Atkinson, E. G., Kanai, M., Martin, A. R., Karczewski, K. J., Ulirsch, J. C., Finucane, H. K., Daly, M. J., Neale, B. M., Santoro, M. L., Koenen, K. C., Maihofer, A. X., Nievergelt, C. M., Okada, Y., & Kamatani, Y. (2020). Tractor: a framework allowing for improved inclusion of admixed individuals in large-scale association studies. *Biorxiv, (2020 05 19)*. https://doi.org/10.1101/2020.05.17.100727

Barnes, J. C., Wright, J. P., Boutwell, B. B., Schwartz, J. A., Connolly, E. J., Nedelec, J. L., & Beaver, K. M. (2014). Demonstrating the validity of twin research in criminology. *Criminology, 52*(4), 588-626. https://doi.org/10.1111/1745-9125.12049

Baron, R. A., & Richardson, D. R. (1994). *Human aggression*. Plenum.

Bartels, M., Derks, E.M., & Boomsma, D.I. (2005). Oorzaken van individuele verschillen in normafwijkend gedrag. Een onderzoek met 7-, 10- en 12-jarige tweelingen. *Tijdschrift voor Criminologie, 47*(2), 129-135.

Bartels, M., Hudziak, J. J., van den Oord, E. J. C. G., van Beijsterveldt, C. E. M., Rietveld, M. J. H., & Boomsma, D. I. (2003). Co-occurrence of aggressive behavior and rule-breaking behavior at age 12: multi-rater analyses. *Behavior Genetics, 33*(5), 607–621. https://doi.org/10.1023/A:1025787019702

Bartels, M., Rietveld, M. J. H., van Baal, G. C. M., & Boomsma, D. I. (2002). Genetic and environmental influences on the development of intelligence. *Behavior Genetics, 32*(4), 237-249. https://doi.org/10.1023/A:1019772628912

Barton, K. (2018). *MuMIn: Multi-Model Inference. R package version 1.40.4.* https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01

Bates, T. C., Maher, B. S., Medland, S. E., McAloney, K., Wright, M. J., Hansell, N. K., Kendler, K. S., Martin, N. G., & Gillespie, N. A. (2018). The nature of nurture: using a virtual-parent design to test parenting effects on children's educational attainment in genotyped families. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies, 21*(2), 73–83. https://doi.org/10.1017/thg.2018.11

Beatton, T., Kidd, M. P., and Machin, S. (2018). Gender crime convergence over twenty years: evidence from Australia. *European Economic Review*, 109, 275-288. https://doi.org/10.1016/j.euroecorev.2018.01.001

Beck, J. J., Hottenga, J. J., Mbarek, H., Finnicum, C. T., Ehli, E. A., Hur, Y. M., Martin, N. G., de, G. E. J. C., Boomsma, D. I., & Davies, G. E. (2019). Genetic similarity assessment of twin-family populations by custom-designed genotyping array. *Twin Research and Human Genetics, 22*(4), 210–219. https://doi.org/10.1017/thg.2019.41

Belsky, D. W., Moffitt, T. E., Sugden, K., Williams, B., Houts, R., McCarthy, J., & Caspi, A. (2013). Development and evaluation of a genetic risk score for obesity. *Biodemography and Social Biology, 59*(1), 85-100. https://doi.org/10.1080/19485565.2013.774628

Berkowitz, L. (1993). *Aggression: Its causes, consequences, and control*. McGraw-Hill.

Besemer, S., Ahmad, S. I., Hinshaw, S. P., & Farrington, D. P. (2017). A systematic review and meta-analysis of the intergenerational transmission of criminal behavior. *Aggression and Violent Behavior, 37*, 161–178. https://doi.org/10.1016/j.avb.2017.10.004

Bijleveld, C. C. & Wijkman, M. (2009). Intergenerational continuity in convictions: a five-generation study. *Criminal Behaviour and Mental Health, 19*(2), 142-155. https://doi.org/10.1002/cbm.714

Björkqvist, K. (1994). Sex differences in physical, verbal, and indirect aggression: a review of recent research. *Sex Roles, 30*(3-4), 177–188. https://doi.org/10.1007/BF01420988

Björkqvist, K. (2018). Gender differences in aggression. *Current Opinion in Psychology, 19*, 39–42. https://doi.org/10.1016/j.copsyc.2017.03.030

Boomsma, D. I. (2015). Aggression in children: Unravelling the interplay of genes and environment through (epi)genetics and metabolomics. *Journal of Pediatric and Neonatal Individualized Medicine*, *4*(2).

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

Boomsma, D. I., Vink, J. M., van Beijsterveldt, C. E. M., de, G. E. J. C., Beem, A. L., Mulder, E. J. C. M., Derks, E. M., Riese, H., Bartels, M., Willemsen, G., van den Berg, M., Kupper, H. M., Polderman, J. C., Posthuma, D., Rietveld, M. J. H., Stubbe, J. H., Knol, L. I., Stroet, T. M., & van Baal, G. C. M. (2002). Netherlands twin register: a focus on longitudinal research. *Twin Research, 5*(5), 401-406. https://doi.org/10.1375/136905202320906174

Boomsma, D. I., de Geus, E. J. C., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., Posthuma, D., van Beijsterveldt, C. E. M., Hudziak, J., Bartels, M., & Willemsen, G. (2006). Netherlands Twin Register: From twins to twin families. *Twin Research and Human Genetics, 9*(6), 849-857. https://doi.org/10.1375/183242706779462426

Boomsma, D. I., Savyuk, V., Hottenga, J. J., Distel, M. A., de Moor, M. H. M., Vink, J. M., Geels, L. M., van Beek, J. H. D. A., Bartels, M., de Geus, E. J. C., & Willemsen, G. (2010). Genetic epidemiology of attention deficit hyperactivity disorder (ADHD index) in adults. *PLoS One, 5*(5). https://doi.org/10.1371/journal.pone.0010621

Boutwell, B. B., & Adams, C. D. (2020). A research note on Mendelian Randomization and causal inference in criminology: Promises and considerations. *Journal of Experimental Criminology*, Article 20200618. https://doi.org/10.1007/s11292-020-09436-9

Brain, P. F., & Susman, E. J. (1997). Hormonal aspects of aggression and violence. In D. M. Stoff, J. Breiling, & J. D. Maser (Eds.). *Handbook of antisocial behavior* (pp. 314-323). John Wiley & Sons.

Branigan, A. R., McCallum, K. J., & Freese, J. (2013). Variation in the heritability of educational attainment: an international meta-analysis. *Social Forces*, *92*(1), 109-140.

Bronfenbrenner, U. (1979). *The ecology of human development.* Harvard University Press.

Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie GÅ, Cho, Y., Howe, L. D., Hughes, A., Boomsma, D. I., Havdahl, A., Hopper, J., Neale, M., Nivard, M. G., Pedersen, N. L., Reynolds, C. A., Tucker-Drob, E. M., Grotzinger, A., Howe, L., ... 23andMe Research Team. (2020). Avoiding dynastic, assortative mating, and population stratification biases in Mendelian Randomization through within-family analyses. *Nature Communications, 11*(1), 3519-3519. https://doi.org/10.1038/s41467-020-17117-4

Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, *35*(2), 138-167. https://doi.org/10.1006/jrpe.2000.2302

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., ... Parkinson, H. (2019). The nhgri-ebi gwas catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research, 47*(1), 1005-1012. https://doi.org/10.1093/nar/gky1120

Burt, C. H., & Simons, R. L. (2014). Pulling back the curtain on heritability studies: Biosocial criminology in the postgenomic era. *Criminology, 52*(2), 223-262. https://doi.org/10.1111/1745-9125.12036

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*(3), 452-459.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Ferguson, L. L., & Gariépy, J. (1989). Growth and aggression: I. Childhood to early adolescence. *Developmental Psychology, 25*(2), 320-330. https://doi.org/10.1037/0012-1649.25.2.320

Cairns, R. B., & Cairns, B. D. (1994). *Life Lines and Risks: Pathways of Youth in Our Time*. Cambridge University Press.

Card, N. A., Sawalani, G. M., Stucky, B. D., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79*(5), 1185-1229.

Carlson, C. S., Matise, T. C., North, K. E., Haiman, C. A., Fesinmeyer, M. D., Buyske, S., Schumacher, F. R., Peters, U., Franceschini, N., Ritchie, M. D., Duggan, D. J., Spencer, K. L., Dumitrescu, L., Eaton, C. B., Thomas, F., Young, A., Carty, C., Heiss, G., Le, M. L., ... PAGE Consortium. (2013). Generalization and dilution of association results from European GWAS in populations of non-European ancestry: The PAGE study. *Plos Biology, 11*(9), Article 1001661. https://doi.org/10.1371/journal.pbio.1001661

Caspi, A., Moffitt, T. E., Morgan, J., Rutter, M., Taylor, A., Arseneault, L., Tully, L., Jacobs, C., Kim-Cohen, J., & Polo-Tomas, M. (2004). Maternal expressed emotion predicts children's antisocial behavior problems: Using monozygotic-twin differences to identify environmental effects on behavioral development. *Developmental Psychology*, *40*(2), 149-161.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach (Monographs in population biology, 16)*. Princeton University Press.

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software, 48*(6): 1-29.

Chen, J., Spracklen, C.N., Marenne, G. et al. (2021). The trans-ancestral genomic architecture of glycemic traits. *Nature Genetics, 53*, 840-860. https://doi-org.vu-nl.idm.oclc.org/10.1038/s41588-021-00852-9

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

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66*(3), 710-722.

Crick, N. R., Werner, N. E., Casas, J. F., O'Brien, K. M., Nelson, D. A., Grotpeter, J. K., & Markon, K. (1999). Childhood aggression and gender: A new look at an old problem. In D. Bernstein (Ed.), *Nebraska symposium on motivation, Vol.45. Gender and motivation* (pp. 75-141). University of Nebraska Press.

Das, S., Forer, L., Schönherr Sebastian, Sidore, C., Locke, A. E., Kwong, A., Vrieze, S. I., Chew, E. Y., Levy, S., McGue, M., Schlessinger, D., Stambolian, D., Loh, P.-R., Iacono, W. G., Swaroop, A., Scott, L. J., Cucca, F., Kronenberg, F., Boehnke, M., … Fuchsberger, C. (2016). Next-generation genotype imputation service and methods. *Nature Genetics, 48*(10), 1284-1287. https://doi.org/10.1038/ng.3656

Demange, A. P., Hottenga, J., Abdellaoui, A., Eilertsen, E. M., Malanchini, M., Domingue, B. W., de Zeeuw, E. L., Rimfeld, K., Eley, T. C., Boomsma, D. I., van Bergen, E., Breen, G., Nivard, M. G., & Cheesman, R. (2021). Estimating effects of parents' cognitive and non-cognitive skills on offspring education using polygenic scores. *Biorxiv.* https://doi.org/*10.1101/2020.09.15.296236*

Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., Baldursson Gísli, Belliveau, R., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Churchhouse, C., Dumont, A., Eriksson, N., Gandal, M., Goldstein, J. I., Grasby, K. L., Grove, J., … Neale, B. M. (2019). Discovery of the first genome-wide significant risk loci for attention-deficit/hyperactivity disorder. *Nature Genetics, 51*(1), 63–75. https://doi.org/10.1038/s41588-018-0269-7

De Zeeuw, E. L., Hottenga, J., Ouwens, K. G., Dolan, C. V., Ehli, E. A., Davies, G. E., Boomsma, D. I., & van Bergen, E. (2020). Intergenerational transmission of education and adhd: effects of parental genotypes. *Behavior Genetics, 50*(4), 221-232. https://doi.org/10.1007/s10519-020-09992-w

De Zeeuw, E. L., Voort, L., Schoonhoven, R., Nivard, M. G., Emery, T., Hottenga, J. J., Willemsen, G. A. H. M., Dykstra, P. A., Zarrabi, N., Kartopawiro, J. D., & Boomsma, D. I. (2021). Safe linkage of cohort and population-based register data in a genomewide association study on health care expenditure. *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies, 24*(2), 103–109. https://doi.org/10.1017/thg.2021.18

Distel, M. A., Rebollo-Mesa, I., Willemsen, G., Derom, C.A., Trull, T.J., Martin, N.G., & Boomsma, D. I. (2009). Familial resemblance of borderline personality disorder features: genetic or cultural transmission? *PLoS One, 4*(4), Article 20090424. https://doi.org/10.1371/journal.pone.0005334.

Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology. 39*(2), 349-371.

Domingue, B. W., Belsky, D. W., Harris, K. M., Smolen, A., McQueen, M. B., Boardman, J. D., & Uddin, M. (2014). Polygenic risk predicts obesity in both white and black young adults. *Plos One, 9*(7). https://doi.org/10.1371/journal.pone.0101596

Domingue, B. W., Belsky, D., Conley, D., Harris, K. M., & Boardman, J. D. (2015). Polygenic influence on educational attainment: new evidence from the national longitudinal study of adolescent to adult health. *Aera Open, 1*(3), 1-13. https://doi.org/10.1177/2332858415599972

Donker, A. G. (2000). Het agressie-gen en andere misverstanden. In: Wetenschappelijk Onderzoek- en Documentatie Centrum WODC (2000). *Biologische factoren van agressief gedrag (Ser. Justitiële verkenningen, 26(3))*. Wetenschappelijk Onderzoek- en Documentatiecentrum, Ministerie van Justitie.

D'Onofrio, B. M., Lahey, B. B., Turkheimer, E., & Lichtenstein, P. (2013). Critical need for family-based, quasi-experimental designs in integrating genetic and social science research. *American Journal of Public Health, 103*(10), 46-55. https://doi.org/10.2105/AJPH.2013.301252.

Dudbridge, F. (2021). Polygenic Mendelian Randomization. *Cold Spring Harbor Perspectives in Medicine, 11*(2), a039586. https://doi.org/10.1101/cshperspect.a039586.

Duffy, D. L. & Martin, N. G. (1994). Inferring the direction of causation in cross-sectional twin data: theoretical and empirical considerations [see comments]. *Genetic Epidemiology*, *11*, 483–502.

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

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum.

Fabio, A., Loeber, R., Balasubramani, G. K., Roth, J., Fu, W., & Farrington, D. P. (2006). Why some generations are more violent than others: Assessment of age, period, and cohort effects. *American Journal of Epidemiology, 164*(2), 151-160.

Farrell, G., Tilley, N., & Tseloni, A. (2014). Why the crime drop? *Crime and Justice, 43*(1), 421-490.

Farrington, D. P. (1989). Early predictors of adolescent aggression and adult violence. *Violence and Victims, 4*(2), 79–100. https://doi.org/10.1891/0886-6708.4.2.79

Farrington, D. P., Gaffney, H., & Ttofi, M. M. (2017). Systematic reviews of explanatory risk factors for violence, offending, and delinquency. *Aggression and Violent Behavior, 33*, 24–36. https://doi.org/10.1016/j.avb.2016.11.004

Ferguson, C. J. (2010). Genetic contributions to antisocial personality and behavior: A meta-analytic review from an evolutionary perspective. *The Journal of Social Psychology*, *150*(2), 160-180. https://doi.org/10.1080/00224540903366503

Frisell, T., Lichtenstein, P., & Långström, N. (2011). Violent crime runs in families: a total population study of 12.5 million individuals. *Psychological medicine, 41*(1), 97-105. https://doi.org/10.1017/S0033291710000462

Frøyland, L. R., & von Soest, T. (2018). Trends in the perpetration of physical aggression among Norwegian adolescents 2007-2015. *Journal of Youth and Adolescence, 47*(9), 1938-1951. https://doi.org/10.1007/s10964-017-0793-2

Gerevich, J., Bácskai, E., & Czobor, P. (2007). The Generalizability of the Buss-Perry Aggression Questionnaire: The generalizability of the AQ. *International Journal of Methods in Psychiatric Research, 16*(3), 124-136. https://doi.org/10.1002/mpr.221

Guerra, N. G., & Leidy, M. S. (2008). Lessons learned: recent advances in understanding and preventing childhood aggression. *Advances in Child Development and Behavior, 36,* 287-330.

Gustavson, D. E., Franz, C. E., Panizzon, M. S., Lyons, M. J., & Kremen, W. S. (2020). Internalizing and externalizing psychopathology in middle age: genetic and environmental architecture and stability of symptoms over 15 to 20 years. *Psychological Medicine, 50*(9), 1530-1538. https://doi.org/10.1017/S0033291719001533

Hagenbeek, F. A., van Dongen, J., Kluft, C., Hankemeier, T., Ligthart, L., Willemsen, G., de Geus, E. J. C., Vink, J. M., Bartels, M., & Boomsma, D. I. (2018). Adult aggressive behavior in humans and biomarkers: a focus on lipids and methylation. *Journal of Pediatric and Neonatal Individualized Medicine, 7*(2), 1-13.

Heath, A. C., Kessler, R. C., Neale, M. C., Hewitt, J. K., Eaves, L. J., & Kendler, K. S. (1993). Testing hypotheses about direction of causation using cross-sectional family data. *Behavior Genetics*, *23*(1), 29–50. https://doi-org.vu-nl.idm.oclc.org/10.1007/BF01067552

Hemani, G., Yang, J., Vinkhuyzen, A., Powell, J. E., Willemsen, G., Hottenga, J. J., Abdellaoui, A., Mangino, M., Valdes, A. M., Medland, S. E., Madden, P. A., Heath, A. C., Henders, A. K., Nyholt, D. R., de, G. E. J., Magnusson, P. K., Ingelsson, E., Montgomery, G. W., Spector, T. D., ... Visscher, P. M. (2013). Inference of the genetic architecture underlying bmi and height with the use of 20,240 sibling pairs. *American Journal of Human Genetics, 93*(5), 865-875. https://doi.org/10.1016/j.ajhg.2013.10.005

Hess, N. H., & Hagen, E. H. (2006). Sex differences in indirect aggression: psychological evidence from young adults. *Evolution and Human Behavior, 27*(3), 231-245. https://doi.org/10.1016/j.evolhumbehav.2005.11.001

Ho, W.-K., Tan, M.-M., Mavaddat, N., Tai, M.-C., Mariapun, S., Li, J., Ho, P.-J., Dennis, J., Tyrer, J. P., Bolla, M. K., Michailidou, K., Wang, Q., Kang, D., Choi, J.-Y., Jamaris, S., Shu, X.-O., Yoon, S.-Y., Park, S. K., Kim, S.-W., ... Antoniou, A. C. (2020). European polygenic risk score for prediction of breast cancer shows similar performance in asian women. *Nature Communications, 11*(1). https://doi.org/10.1038/s41467-020-17680-w

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. Biometrical Journal. *Biometrische Zeitschrift, 50*(3), 346-363. https://doi.org/10.1002/bimj.200810425

Huesmann, L. R., Dubow, E. F., & Boxer, P. (2009). Continuity of aggression from childhood to early adulthood as a predictor of life outcomes: implications for the adolescent-limited and life-course-persistent models. *Aggressive Behavior, 35*(2), 136–49. https://doi.org/10.1002/ab.20300

Huppertz, C., Bartels, M., de Zeeuw, E. L., van Beijsterveldt, C. E. M., Hudziak, J. J., Willemsen, G., Boomsma, D. I., & de Geus, E. J. C. (2016). Individual differences in exercise behavior: stability and change in genetic and environmental determinants from age 7 to 18. *Behavior Genetics, 46*(5), 665-679. https://doi.org/10.1007/s10519-016-9799-x

Ip, H. F., van der Laan, C. M., Brikell, I., Sánchez-Mora, C., Nolte, I. M., St. Pourcain, B., Bolhuis, K., Palviainen, T., Zafarmand, H., Colodro-Conde, L., Gordon, S., Zayats, T., Aliev, F., Jiang, C., Wang, C. A., Saunders, G., Karhunen, V., Hammerschlag, A. R., ... Boomsma, D. I. (2021). Genetic association study of childhood aggression across raters, instruments, and age. *Translational Psychiatry, 11*, Article 413. https://doi.org/10.1038/s41398-021-01480-x

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Turner, L. V., Ahmeti-Pronaj, A., Au, A., Maese, C. A., Bellina, M., Caldas, J. C., Chen, Y.-C., Csemy, L., da Rocha, M. M., Decoster, J., Dobrean, A., Ezpeleta, L., Fontaine, J. R. J., Funabiki, Y., Guðmundsson Halldór S, Harder, V. S., ... Zasepa, E. (2015). Syndromes of self-reported psychopathology for ages 18-59 in 29 societies. *Journal of Psychopathology and Behavioral Assessment, 37*(2), 171-183. https://doi.org/10.1007/s10862-014-9448-8

Ivanova, M. Y., Dobrean, A., Dopfner, M., Erol, N., Fombonne, E., Fonseca, A. C., Frigerio, A., Grietens, H., Hannesdottir, H., Kanbayashi, Y., Lambert, M., Achenbach, T. M., Larsson, B., Leung, P., Liu, X., Minaei, A., Mulatu, M. S., Novik, T. S., Oh, K. J., ... Chen, W. J. (2007). Testing the 8-syndrome structure of the child behavior checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology, 36*(3), 405-417.

Junger, M., Greene, J., Schipper, R., Hesper, F., & Estourgie, V. (2013). Parental criminality, family violence and intergenerational transmission of crime within a birth cohort. *European Journal on Criminal Policy and Research, 19*(2), 117-133. https://doi.org/10.1007/s10610-012-9193-z

Kann, L., McManus, T., Harris, W. A., Shanklin, S. L., Flint, K. H., Hawkins, J., Queen, B., Lowry, R., Olsen, E. O. M., Chyen, D., Whittle, L., Thornton, J., Lim, C., Yamakawa, Y., Brener, N., & Zaza, S. (2016). Youth risk behavior surveillance. *MMWR. Surveillance Summaries, 65*(6), 1-174. https://doi.org/10.15585/mmwr.ss6506a1

Karriker-Jaffe, K. J., Foshee, V. A., Ennett, S. T., & Suchindran, C. (2008). The development of aggression during adolescence: sex differences in trajectories of physical and social aggression among youth in rural areas. *Journal of Abnormal Child Psychology, 36*(8), 1227-1236. https://doi.org/10.1007/s10802-008-9245-5

Keller, M. C., Medland, S. E., Duncan, L. E., Hatemi, P. K., Neale, M. C., Maes, H. M. M., & Eaves, L. J. (2009). Modeling extended twin family data I: Description of the cascade model. *Twin Research & Human Genetics, 12*(1), 8-18. https://doi.org/10.1375/twin.12.1.8

Kendler, K. S., & Eaves, L. J. (1986). Models for the joint effects of genotype and environment on liability to psychiatric illness. *American Journal of Psychiatry, 143*(3), 279-289.

Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., Benonisdottir, S., Oddsson, A., Halldorsson, B. V., Masson, G., Gudbjartsson, D. F., Helgason, A., Bjornsdottir, G., Thorsteinsdottir, U., & Stefansson, K. (2018). The nature of nurture: Effects of parental genotypes. *Science*, *359*(6374), 424-428. https://doi.org/10.1126/science.aan6877

Labella, M. H., & Masten, A. S. (2018). Family influences on the development of aggression and violence. *Current Opinion in Psychology, 19*, 11–16. https://doi.org/10.1016/j.copsyc.2017.03.028

Lamb, D. J., Middeldorp, C. M., van Beijsterveldt, C. E. M., Bartels, M., van der Aa, N., Polderman, T. J. C., & Boomsma, D. I. (2010). Heritability of anxious-depressive and withdrawn behavior: Age-related changes during adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(3), 248-255. https://doi.org/10.1016/j.jaac.2009.11.014

Latvala, A., Kuja-Halkola, R., Rück, C., D'Onofrio, B. M., Jernberg, T., Almqvist, C., Mataix-Cols, D., Larsson, H., & Lichtenstein, P. (2016). Association of resting heart rate and blood pressure in late adolescence with subsequent mental disorders: A longitudinal population study of more than 1 million men in Sweden. *Jama Psychiatry, 73*(12), 1268-1275. https://doi.org/10.1001/jamapsychiatry.2016.2717

Laub, J. H., & Sampson, R. J. (2006). *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70.* Harvard University Press.

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

Ligthart, L., van Beijsterveldt, C. E. M., Kevenaar, S. T., de Zeeuw, E., van Bergen, E., Bruins, S., Pool, R., Helmer, Q., van, D. J., Hottenga, J. J., Van't, E. D., Dolan, C. V., Davies, G. E., Ehli, E. A., Bartels, M., Willemsen, G., de, G. E. J. C., & Boomsma, D. I. (2019). The Netherlands Twin register: Longitudinal research based on twin and twin-family designs. *Twin Research and Human Genetics, 22*(6), 623-636. https://doi.org/10.1017/thg.2019.93

Linner, R.K., Koellinger, P.D., Mallard, T.T., Madole, J.W., Grotzinger, A.D., Tucker-Drob, E.M. et al. (2021). Multivariate genomic analysis of 1.5 million people identifies genes related to addiction, antisocial behavior, and health. *Nature Neuroscience, 24*(10), 1367-1376. https://doi.org/10.1038/s41593-021-00908-3.

Liu, J., Lewis, G., & Evans, L. (2013). Understanding aggressive behaviour across the lifespan. *Journal of Psychiatric and Mental Health Nursing, 20*(2), 156-168. https://doi.org/10.1111/j.1365-2850.2012.01902.x

Loeber, R., & Stouthamer-Loeber, M. (1998). Development of juvenile aggression and violence. some common misconceptions and controversies. *The American Psychologist, 53*(2), 242-259.

Lorenz, K. (1966). *Das sogenannte böse: Zur naturgeschichte der aggression*. Borotha-Schoeler.

Lloyd-Jones, L. R., Zeng, J., Sidorenko, J., Yengo Loïc, Moser, G., Kemper, K. E., Wang, H., Zheng, Z., Magi, R., Esko Tõnu, Metspalu, A., Wray, N. R., Goddard, M. E., Yang, J., & Visscher, P. M. (2019). Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nature Communications, 10*(1), 1–11. https://doi.org/10.1038/s41467-019-12653-0

Lubke, G. H., Miller, P. J., Verhulst, B., Bartels, M., van Beijsterveldt, T., Willemsen, G., Boomsma, D. I., & Middeldorp, C. M. (2016). A powerful phenotype for gene-finding studies derived from trajectory analyses of symptoms of anxiety and depression between age 7 and 18. *American Journal of Medical Genetics, 171*(7), 948-957. https://doi.org/10.1002/ajmg.b.32375

Lupton, M. K., Robinson, G. A., Adam, R. J., Rose, S., Byrne, G. J., Salvado, O., Pachana, N. A., Almeida, O. P., McAloney, K., Gordon, S. D., Raniga, P., Fazlollahi, A., Xia, Y., Ceslis, A., Sonkusare, S., Zhang, Q., Kholghi, M., Karunanithi, M., Mosley, P. E., ... Breakspear, M. (2021). A prospective cohort study of prodromal Alzheimer's disease: Prospective imaging study of ageing: genes, brain and behaviour (pisa). *Neuroimage: Clinical, 29*, 102527–102527. https://doi.org/10.1016/j.nicl.2020.102527

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *The American Psychologist, 45*(4), 513-520.

Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W. M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics, 26*(22), 2867-2873. https://doi.org/10.1093/bioinformatics/btq559

Margolin, G., Ramos, M. C., Timmons, A. C., Miller, K. F., & Han, S. C. (2016). Intergenerational transmission of aggression: Physiological regulatory processes. *Child Development Perspectives, 10*(1), 15-21. https://doi.org/10.1111/cdep.12156

Mark, K. M., Pike, A., Latham, R. M., & Oliver, B. R. (2017). Using twins to better understand sibling relationships. *Behavior Genetics, 47*(2), 202-214. https://doi.org/10.1007/s10519-016-9825-z

Martin, N., Boomsma, D.I, & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics, 17*, 387-392.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics, 51*(4), 584–591. https://doi.org/10.1038/s41588-019-0379-x

Matthews, B., & Minton, J. (2017). Rethinking one of criminology's 'brute facts': The age-crime curve and the crime drop in Scotland. *European Journal of Criminology, 15*(3), 296-320. https://doi.org/10.1177/1477370817731706

Maxwell, J. P. (2007). Development and preliminary validation of a Chinese version of the Buss-Perry aggression questionnaire in a population of Hong Kong Chinese. *Journal of Personality Assessment, 88*(3), 284-294.

Mednick, S. A., Gabrielli, W. F., & Hutchings, B. (1984). Genetic influences in criminal convictions. Evidence from an adoption cohort. *Science, 224*(4651), 891-894.

Meyer-Lindenberg, A., Buckholtz, J. W., Kolachana, B., R. Hariri, A., Pezawas, L., Blasi, G., Wabnitz, A., Honea, R., Verchinski, B., Callicott, J. H., Egan, M., Mattay, V., & Weinberger, D. R. (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences of the United States of America, 103*(16), 6269-6274. https://doi.org/10.1073/pnas.0511311103

Mills, M. C., & Rahal, C. (2020). The GWAS diversity monitor tracks diversity by disease in real time. *Nature Genetics, 52*(3), 242-243. https://doi.org/10.1038/s41588-020-0580-y

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. Routledge. https://doi.org/10.4324/9780203821961

Minicaˇ, C. C., Dolan, C. V., Kampert, M. M., Boomsma, D. I., and Vink, J. M. (2015). Sandwich corrected standard errors in family-based genome-wide association studies. *European Journal of Human Genetics, 23*(3), 388-394. https://doi.org/10.1038/ejhg.2014.94

Mitchell, B. L., Campos, A. I., Renteria, M. E., Parker, R., Sullivan, L., Mcaloney, K., Medland, S. E., Lind, P. A., Martin, N. G., Couvy-Duchesne, B., Gillespie, N. A., Scott, J., Hickie, I. B., & Zietsch, B. P. (2019). Twenty-five and up (25up) study: A new wave of the brisbane longitudinal twin study. *Twin Research and Human Genetics*, *22*(3), 154-163. https://doi.org/10.1017/thg.2019.27

Mitchell, B. L., Kirk, K. M., McAloney, K., Scott, J. G., Gillespie, N. A., Medland, S., Martin, N. G., Wright, M. J., McGrath, J. J., O'Callaghan, V. S., Davenport, T. A., Carpenter, J. S., Hickie, I. B., Hermens, D. F., Christensen, H., & Burns, J. M. (2021). 16UP: Outline of a study investigating wellbeing and information and communication technology use in adolescent twins. *Twin Research and Human Genetics*, *23(6), 345-357*. https://doi.org/10.1017/thg.2020.83

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*(4), 674-701. https://doi.org/10.1037//0033-295X.100.4.674

Morosoli, J.J. (2020). Overcoming genetic determinism: A psychological perspective on public understanding of human complex trait genetics. (Unpublished doctoral dissertation). University of Queensland, Brisbane, Australia.

Murray, J., & Farrington, D. P. (2010). Risk factors for conduct disorder and delinquency: Key findings from longitudinal studies. *Canadian Journal of Psychiatry, 55*(10), 633-642.

Neale, M. C., Hunter, M. D., Pritkin, J., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). Openmx 2.0: extended structural equation and statistical modeling. *Psychometrika, 81*(2), 535-549. https://doi.org/10.1007/s11336-014-9435-8

Neugebauer, R. Hoek, H. W., & Susser, E. (1999). Prenatal exposure to wartime famine and development of antisocial personality disorder in early adulthood. *Jama, 282*(5), 455-462.

Nieuwenhuis, J. G., van Ham, M., Yu, R., Branje, S., Meeus, W., & Hooimeijer, P. (2017). Being poorer than the rest of the neighborhood: Relative deprivation and problem behavior of youth. *Journal of Youth and Adolescence*, *46*(9), 1891-1904. https://doi.org/10.1007/s10964-017-0668-6

Odgers, C. L., Donley, S., Caspi, A., Bates, C. J., & Moffitt, T. E. (2015). Living alongside more affluent neighbors predicts greater involvement in antisocial behavior among low-income boys. *Journal of Child Psychology and Psychiatry, 56*(10), 1055-1064. https://doi.org/10.1111/jcpp.12380

Odintsova, V. V., Roetman, P. J., Ip, H. F., Pool, R., van der Laan, C. M., Tona, K. D., Vermeiren, R. R. J. M., & Boomsma, D. I. (2019). Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool. *Psychiatric Genetics, 29*(5), 170-190. https://doi.org/10.1097/YPG.0000000000000239

Pappa, I., St Pourcain, B., Benke, K., Cavadino, A., Hakulinen, C., Nivard, M. G., Nolte, I. M., Tiesler, C. M. T., Bakermans-Kranenburg, M. J., Davies, G. E., Evans, D. M., Geoffroy, M.-C., Grallert, H., Groen-Blokhuis, M. M., Hudziak, J. J., Kemp, J. P., Keltikangas-Jarvinen, L., McMahon, G., Mileva-Seitz, V. R., … Tiemeier, H. (2016). A genome-wide approach to children's aggressive behavior: The Eagle consortium. *American Journal of Medical Genetics. Part B: Neuropsychiatric Genetics, 171*(5), 562-572. https://doi.org/10.1002/ajmg.b.32333

Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic Epidemiology, 32*(4), 381-385. https://doi.org/10.1002/gepi.20303

Pickett, W., Molcho, M., Elgar, F. J., Brooks, F., de, L. M., Rathmann, K., ter, B. T. F., Nic, G. S., Sigmundová D, Gaspar, de M. M., Craig, W., Walsh, S. D., Harel-Fisch, Y., & Currie, C. (2013). Trends and socioeconomic correlates of adolescent physical fighting in 30 countries. *Pediatrics, 131*(1), 18-26. https://doi.org/10.1542/peds.2012-1614

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2017). *NLME: Linear and Nonlinear Mixed Effects Models. R package version 3.1-131*. https://CRAN.R-project.org/package=nlme

Plomin, R. (2014). Genotype-environment correlation in the era of DNA. *Behavior Genetics, 44*(6), 629-638. https://doi.org/10.1007/s10519-014-9673-7

Plomin, R., & Daniels, D. (2011). Why are children in the same family so different from one another? *International Journal of Epidemiology, 40*(3), 563-582. https://doi.org/10.1093/ije/dyq148

Plomin, R., DeFries, J., & Fulker, D. (1977). Genotype–environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*, 309–322.

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), 702-709. https://doi.org/10.1038/ng.3285

Porsch, R. M., Middeldorp, C. M., Cherny, S. S., Krapohl, E., van Beijsterveldt, C. E. M., Loukola, A., Korhonen, T., Pulkkinen, L., Corley, R., Rhee, S., Kaprio, J., Rose, R. R., Hewitt, J. K., Sham, P., Plomin, R., Boomsma, D. I., & Bartels, M. (2016). Longitudinal heritability of childhood aggression. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 171*(5), 697-707. https://doi.org/10.1002/ajmg.b.32420

Pulkkinen, L., & Pitkänen Tuuli. (1993). Continuities in aggressive behavior from childhood to adulthood. *Aggressive Behavior, 19*(4), 249-263. https://doi.org/10.1002/1098-2337

Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research, 5*(6), 554-571. https://doi.org/10.1375/136905202762342026

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., … Sham, P.C. (2007). PLINK: A toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics, 81*(3), 559-575

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raine, A. (2008). From genes to brain to antisocial behavior. *Current Directions in Psychological Science, 17*(5), 323-328. https://doi.org/10.1111/j.1467-8721.2008.00599.x

Raine, A., Fung, A. L., Portnoy, J., Choy, O., & Spring, V. L. (2014). Low heart rate as a risk factor for child and adolescent proactive aggressive and impulsive psychopathic behavior. *Aggressive Behavior, 40*(4), 290-299. https://doi.org/10.1002/ab.21523

Ramirez, J. M. (2003). Hormones and aggression in childhood and adolescence. *Aggression and Violent Behavior, 8*(6), 621-644. https://doi.org/10.1016/S1359-1789(02)00102-7

Repetti, R., Taylor, S., & Seeman, T. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin, 128*(2), 330-366.

Rhee, S. H., & Waldman, I. D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin*, *128*(3), 490-529.

Rietveld, M. J. H., van der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E. & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research and Human Genetics, 3*(3), 134-141. https://doi.org/10.1375/twin.3.3.134

Rosenberg, N. A. (2021). A population-genetic perspective on the similarities and differences among worldwide human populations. *Human Biology, 92*(3), 135-152. https://doi.org/10.13110/humanbiology.92.3.02

Rosser, Z. H., Zerjal, T., Hurles, M. E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., Beckman, G., Beckman, L., Bertranpetit, J., Bosch, E., Bradley, D. G., Brede, G., Cooper, G., Côrte-Real HB, de, K. P., ... Jobling, M. A. (2000). Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *American Journal of Human Genetics, 67*(6), 1526-1543.

Rud, I., van Klaveren, C., Groot, W., & Maassen van den Brink, H. (2018). What drives the relationship between early criminal involvement and school dropout? *Journal of Quantitative Criminology*, *34*(1), 139–166. https://doi.org/10.1007/s10940-016-9326-5

Sampson, R. J., & Laub, J. H. (1994). Urban poverty and the family context of delinquency: A new look at structure and process in a classic study. *Child Development, 65*(2), 523-540.

Sariaslan, A., Långström N, D'Onofrio, B., Hallqvist, J., Franck, J., & Lichtenstein, P. (2013). The impact of neighbourhood deprivation on adolescent violent criminality and substance misuse: A longitudinal, quasi-experimental study of the total Swedish population. *International Journal of Epidemiology*, *42*(4), 1057-1066. https://doi.org/10.1093/ije/dyt066

Selzam, S., Ritchie, S. J., Pingault, J. B., Reynolds, C. A., O'Reilly, P. F., & Plomin, R. (2019). Comparing within- and between-family polygenic score prediction. *American Journal of Human Genetics, 105*(2), 351-363. https://doi.org/10.1016/j.ajhg.2019.06.006

Smith, G. D., & Ebrahim, S. (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International journal of epidemiology, 32*(1), 1–22.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE.

Spivack, G., & Spotts, J. (1966). *The Devereux Child Behavior (DCB) Rating Scale*. The Devereux Foundation.

Steinberg, L. D. (1986). Latchkey children and susceptibility to peer pressure: an ecological analysis. *Developmental Psychology, 22*(4), 433-439.

Sullivan, P. F., Montgomery, G. W., Hottenga, J. J., Wray, N. R., Boomsma, D. I., & Martin, N. G. (2006). Empirical evaluation of the genetic similarity of samples from twin registries in Australia and The Netherlands using 359 STRP markers. *Twin Research and Human Genetics, 9*(4), 600-602. https://doi.org/10.1375/183242706778025026

Swagerman, S. C., van Bergen, E., Dolan, C., de Geus, E. J. C., Koenis, M. M. G., Hulshoff Pol, H. E., & Boomsma, D. I. (2017). Genetic transmission of reading ability. *Brain and Language, 172*, 3-8. https://doi.org/10.1016/j.bandl.2015.07.008

Thomson, N. D., Bozgunov, K., Psederska, E., & Vassileva, J. (2019). Sex differences on the four-facet model of psychopathy predict physical, verbal, and indirect aggression. *Aggressive Behavior, 45*(3), 265-274. https://doi.org/10.1002/ab.21816

Tielbeek, J. J., Johansson, A., Polderman, T. J. C., Rautiainen, M.-R., Jansen, P., Taylor, M., Tong, X., Lu, Q., Burt, A. S., Tiemeier, H., Viding, E., Plomin, R., Martin, N. G., Heath, A. C., Madden, P. A. F., Montgomery, G., Beaver, K. M., Waldman, I., Gelernter, J., … Posthuma, D. (2017). Genome-wide association studies of a broad spectrum of antisocial behavior. *Jama Psychiatry, 74*(12), 1242–1250. https://doi.org/10.1001/jamapsychiatry.2017.4547

Tielbeek, J. J., Vink, J. M., Polderman, T. J., Popma, A., Posthuma, D., & Verweij, K. J. (2018). Genetic correlation of antisocial behaviour with alcohol, nicotine, and cannabis use. *Drug and alcohol dependence*, *187*, 296-299.

Tolan, P. H., Dodge, K., & Rutter, M. (2013). Tracking the multiple pathways of parent and family influence on disruptive behavior disorders. In Tolan, P. H., and Leventhal, B. (Eds). *Disruptive behavior disorders* (pp. 161-191). Springer.

Tremblay, R. E. (2010). Developmental origins of disruptive behaviour problems: The 'original sin' hypothesis, epigenetics and their consequences for prevention. *Journal of Child Psychology and Psychiatry, 51*(4), 341-367. https://doi.org/10.1111/j.1469-7610.2010.02211.x

Tremblay, R. E., Nagin, D. S., Séguin, J. R, Zoccolillo, M., Zelazo, P. D., Boivin, M., Pérusse Daniel, & Japel, C. (2004). Physical aggression during early childhood: trajectories and predictors. *Pediatrics, 114*(1), 43-50. https://doi.org/10.1542/peds.114.1.e43

Tuvblad, C., & Baker, L. A. (2011). Human aggression across the lifespan: genetic propensities and environmental moderators. Advances in Genetics, 75, 171–214. https://doi.org/10.1016/B978-0-12-380858-5.00007-1

Underwood, M. K. (2003). *Social Aggression Among Girls*. Guilford.

Underwood, M. K., Beron, K. J., Gentsch, J. K., Galperin, M. B., and Risser, S. D. (2008). Family correlates of children's social and physical aggression with peers: Negative interparental conflict strategies and parenting styles. *International Journal of Behavioral Development, 32*(6), 549–562. https://doi.org/10.1177/0165025408097134

United States Department of Justice, Federal Bureau of Investigation (2019). *Crime in the United States, 2019*. Retrieved 22 March 2021, from https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019.

Valdivia-Peralta, M., Fonseca-Pedrero, E., González-Bravo, L., & Lemos-Giráldez, S. (2014). Psychometric properties of the AQ aggression scale in Chilean students. *Psicothema*, *26*(1), 39-46. https://doi.org/10.7334/psicothema2013.84

Van Beek, J. A. van, Moor, M. H. M. de, Geels, L. M., Willemsen, G. & Boomsma, D.I. (2014). Explaining individual differences in alcohol intake in adults. Evidence for genetic and cultural transmission? *Journal of Studies on Alcohol and Drugs, 75*(2), 201-210.

Van Beijsterveldt, C. E. M., Groen-Blokhuis, M., Hottenga, J. J., Franič Sanja, Hudziak, J. J., Lamb, D., Huppertz, C., de Zeeuw, E., Nivard, M., Schutte, N., Swagerman, S., Glasner, T., van Fulpen, M., Brouwer, C., Stroet, T., Nowotny, D., Ehli, E. A., Davies, G. E., Scheet, P., ... 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(1), 252–267. https://doi.org/10.1017/thg.2012.118

Van Beijsterveldt, C. E. M., Verhulst, F. C., Molenaar, P. C. M., & Boomsma, D. I. (2004). The genetic basis of problem behavior in 5-year-old Dutch twin pairs. *Behavior Genetics, 34*(3), 229–242. https://doi.org/10.1023/B:BEGE.0000017869.30151.fd

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: multivariate imputation by chained equations in r. *Journal of Statistical Software*, *45*(3), 1–67.

Van de Weijer, S. G. A. Bijleveld, C. C. J. H., and Blokland, A. A. J. (2014). The intergenerational transmission of violent offending. *Journal of Family Violence, 29*(2), 109-118. https://doi.org/10.1007/s10896-013-9565-2

Van der Laan, A.M., & Goudriaan, H. (2016). Monitor Jeugdcriminaliteit. Ontwikkelingen in de jeugdcriminaliteit 1997 tot 2015. *Cahier 2016-1*. www.wodc.nl/binaries/cahier-2016-1-volledige-tekst_tcm28-74162.pdf

Van der Laan, C. M., Morosoli-Garcia, J. J., van de Weijer, S. G. A., Colodro-Conde, L., ACTION consortium, Lupton, M. K., ... Boomsma, D. I. (2021-c). Continuity of genetic risk for aggressive behavior across the life-course. *Behavior Genetics, 51*, 592-606. https://doi.org/10.1007/s10519-021-10076-6

Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2019). Een tweelingstudie naar indicatoren van genetische en culturele transmissie. *Tijdschrift Voor Criminologie*, *61*(2), 148-164. https://doi.org/10.5553/TvC/0165182X2019061002003

Van der Laan, C. M., van de Weijer, S. G. A, Nivard, M. G., & Boomsma, D. I. (2021-b). Familial Clustering of Trends in Aggression. *Journal of Quantitative Criminology. https://doi.org/10.1007/s10940-021-09523-8*

Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2021-a). Genetische onderzoekdesigns in de criminologie: Een 'toolbox' voor onderzoek naar causaliteit en intergenerationele transmissie. *Tijdschrift Voor Criminologie*, *63*(3), 347-358. https://doi.org/10.5553/TvC/0165182X2021063003004

Van de Rakt, M., Nieuwbeerta, P., & De Graaf, N. D. (2008). Like father, like son: The relationships between conviction trajectories of fathers and their sons and daughters. *The British journal of criminology, 48*(4), 538-556. https://doi.org/10.1093/bjc/azn014

Van Dongen, J., Nivard, M. G., Baselmans, B. M. L., Zilhão Nuno R, Ligthart, L., Heijmans, B. T., Bartels, M., Boomsma, D. I., Craig, J., & Saffery, R. (2015). Epigenome-wide association study of aggressive behavior. *Twin Research and Human Genetics, 18*(6), 686-698. https://doi.org/10.1017/thg.2015.74

Vassos, E., Di, F. M., Coleman, J., Iyegbe, C., Prata, D., Euesden, J., O'Reilly, P., Curtis, C., Kolliakou, A., Patel, H., Newhouse, S., Traylor, M., Ajnakina, O., Mondelli, V., Marques, T. R., Gardner-Sood, P., Aitchison, K. J., Powell, J., Atakan, Z., … Breen, G. (2017). An examination of polygenic score risk prediction in individuals with first-episode psychosis. *Biological Psychiatry*, *81*(6), 470-477. https://doi.org/10.1016/j.biopsych.2016.06.028

Verhulst, F., Achenbach, T., & Ende, J. van der (2013). *Handleiding ASEBA: vragenlijsten voor leeftijden 6 t/m 18 jaar: Cbcl/6-18 – ysr – trf.* ASEBA Nederland.

Veroude, K., Zhang-James, Y., Fernàndez-Castillo, N., Bakker, M. J., Cormand, B., & Faraone, S.V. (2016). Genetics of aggressive behavior: an overview. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 171*(1), 3-43. https://doi.org/10.1002/ajmg.b.32364

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. https://doi.org/10.18637/jss.v036.i03

Vilhjálmsson Bjarni J, Yang, J., Finucane, H. K., Gusev, A., Lindström Sara, Ripke, S., Genovese, G., Loh, P.-R., Bhatia, G., Do, R., Hayeck, T., Won, H.-H., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., … Price, A. L. (2015). Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *The American Journal of Human Genetics*, *97*(4), 576-592. https://doi.org/10.1016/j.ajhg.2015.09.001

Vinkhuyzen, A. A. E., Van Der Sluis, S., de Geus, E. J. C., Boomsma, D. I., & Posthuma, D. (2010). Genetic influences on 'environmental' factors. *Genes, Brain and Behavior, 9*(3), 276-287. https://doi.org/10.1111/j.1601-183X.2009.00554.x

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*, *90*(1), 7-24. https://doi.org/10.1016/j.ajhg.2011.11.029

Visscher, P. M., Medland, S. E., Ferreira, M. A. R., Morley, K. I., Zhu, G., Cornes, B. K., Montgomery, G. W., Martin, N. G., & Mackay, T. (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *Plos Genetics, 2*(3), Article 20060324. https://doi.org/10.1371/journal.pgen.0020041

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: biology, function, and translation. *The American Journal of Human Genetics*, *101*(1), 5-22. https://doi.org/10.1016/j.ajhg.2017.06.005

Vuoksimaa, E., Rose, R. J., Pulkkinen, L., Palviainen, T., Rimfeld, K., Lundström Sebastian, Bartels, M., Beijsterveldt, C., Hendriks, A., Zeeuw, E. L., Plomin, R., Lichtenstein, P., Boomsma, D. I., & Kaprio, J. (2021). Higher aggression is related to poorer academic performance in compulsory education. *Journal of Child Psychology and Psychiatry, 62*(3), 327-338. https://doi.org/10.1111/jcpp.13273

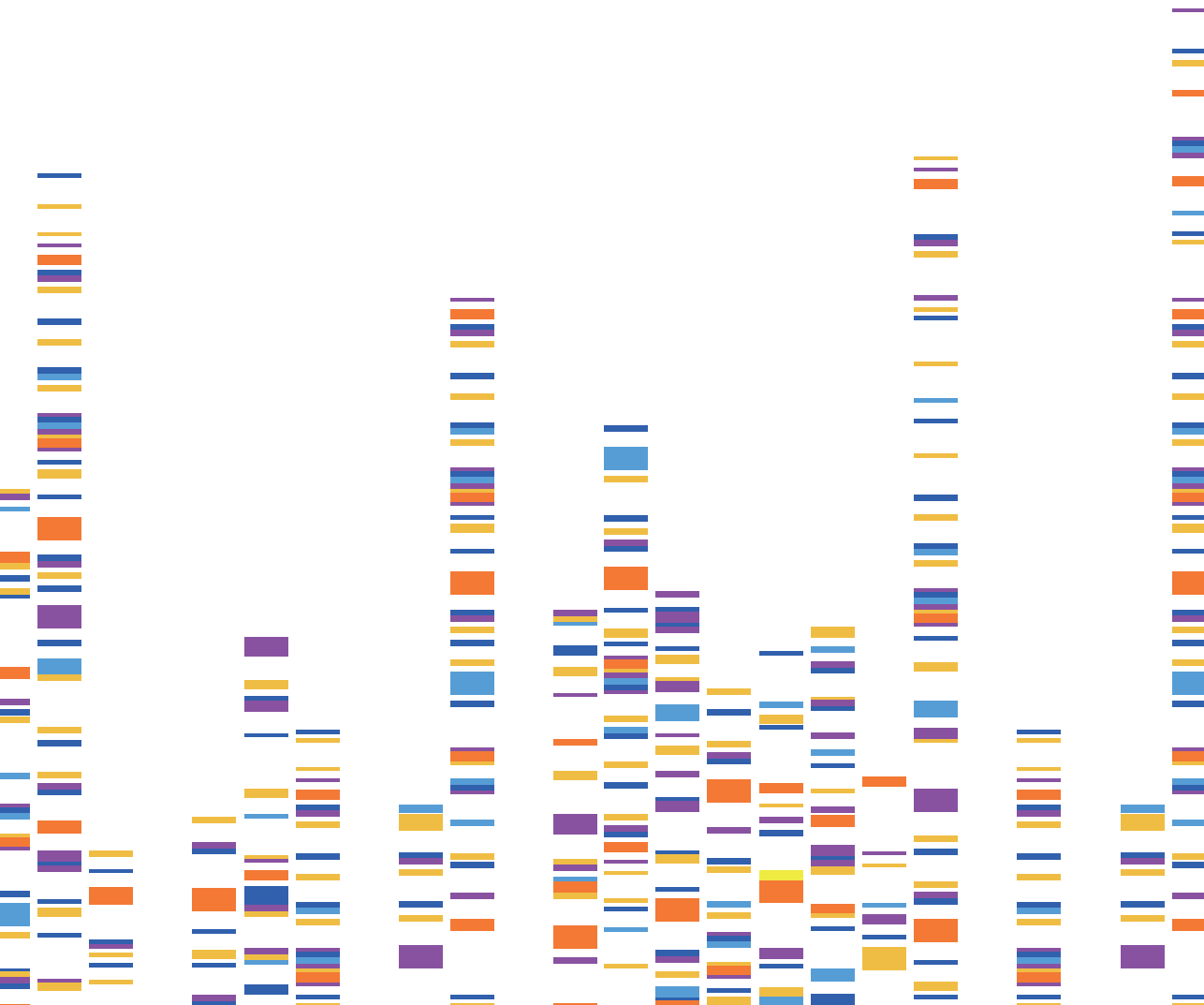
Waltes, R., Chiocchetti, A. G., & Freitag, C. M. (2016). The neurobiological basis of human aggression: A review on genetic and epigenetic mechanisms. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics, 171*(5), 650-675. https://doi.org/10.1002/ajmg.b.32388

Ware, E. B., Schmitz, L. L., Faul, J., Mitchell, C., Smith, J. A., Weir, D., Gard, A., Zhao, W., & Kardia, S. L. R. (2017). Heterogeneity in polygenic scores for common human traits. *Biorxiv*, *(2017 02 05)*. https://doi.org/10.1101/106062

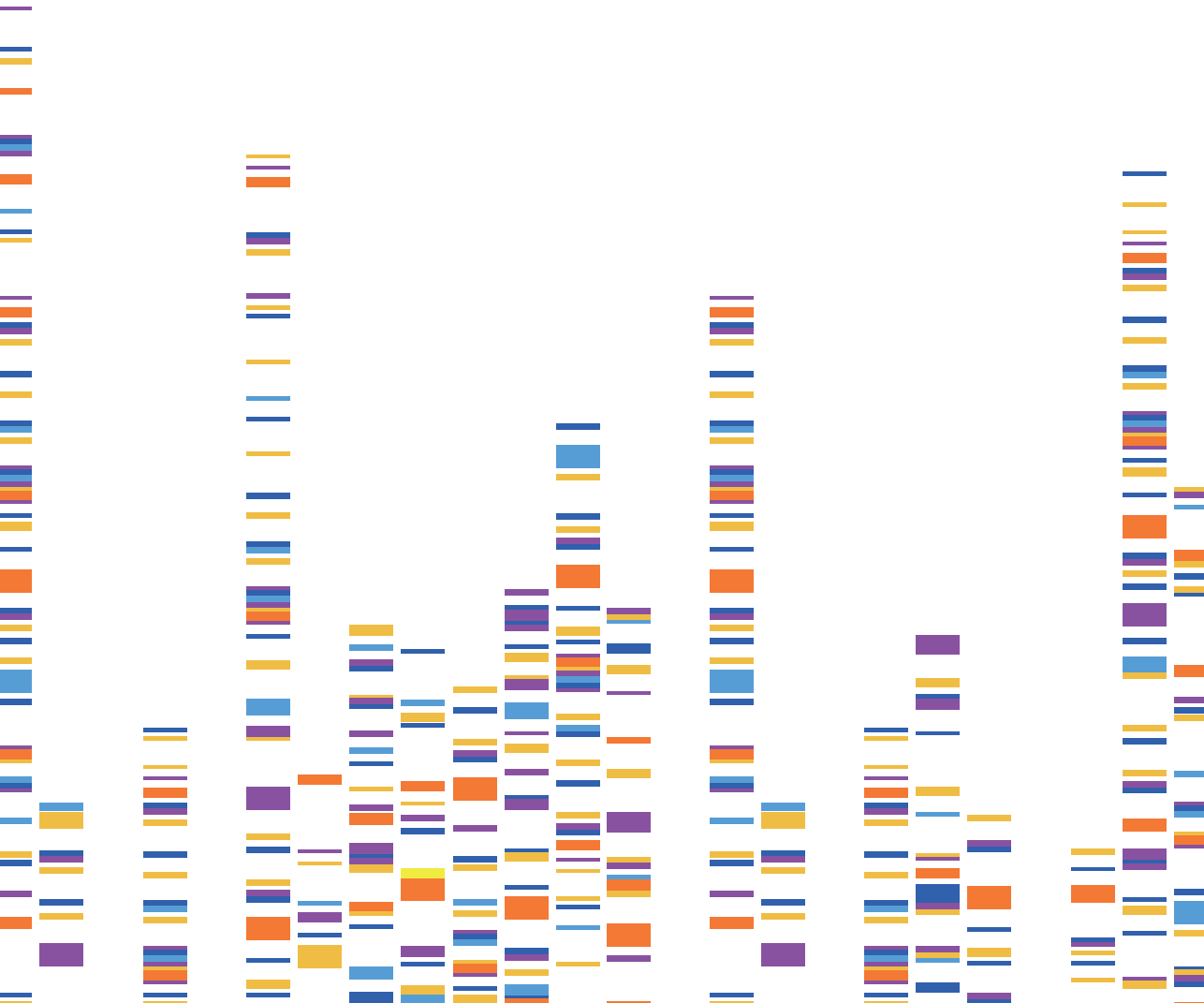
Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Willems, Y. E., de Zeeuw, E. L., van Beijsterveldt, C., Boomsma, D. I., Bartels, M., & Finkenauer, C. (2020). Out of control: Examining the association between family conflict and self-control in adolescence in a genetically sensitive design. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(2), 254-262. https://doi.org/10.1016/j.jaac.2019.02.017

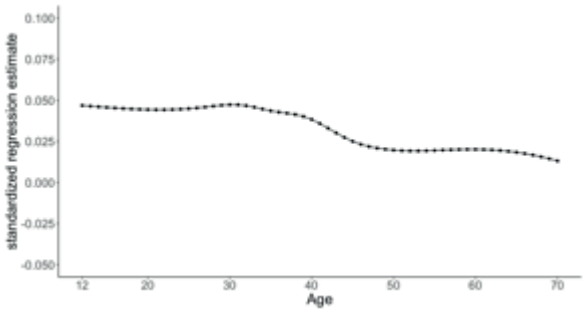
Yu, J. J., & Gamble, W. C. (2008). Familial correlates of overt and relational aggression between young adolescent siblings. *Journal of Youth and Adolescence, 37*(6), 655-673.



10
Supplements



A. Wide weights distribution



B. Narrow weights distribution

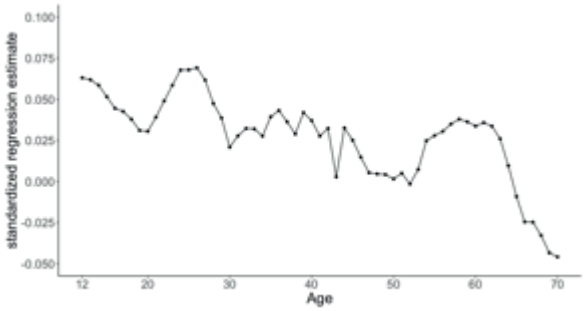


Figure S4.1. The Netherlands: Regression estimates for PGS p< 0.1 with two different weight distributions for each age.

Note: On the left: results from models with a wider distribution of weights, with low sensitivity to specific age effects. On the right: results from models with a narrower distribution of weights, with high sensitivity to specific age effects. See Figure S4.2 for an example of the differences in weight distributions.

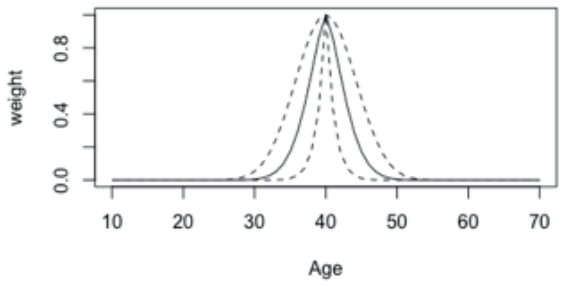
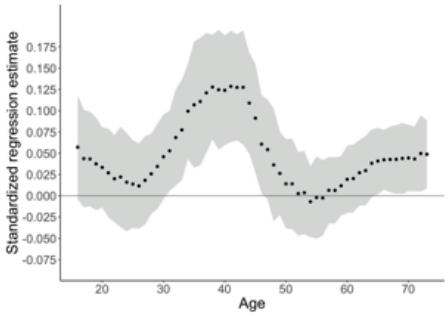


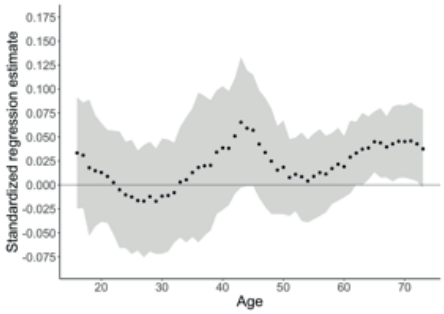
Figure S4.2. Example of wider and narrower distributions at age 40, used for Supplements Figure S4.1

Note: The solid line reflects the distribution of weights used in our main analyses. The dotted lines represent the wider and narrower distributions.

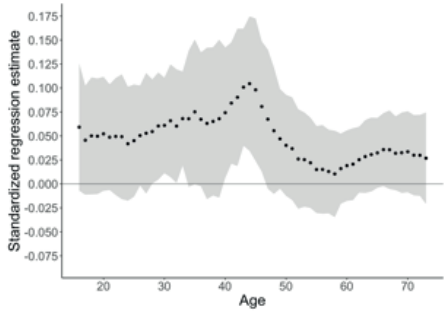
A. Physical Aggression



B. Hostility



C. Anger



D. Verbal Aggression

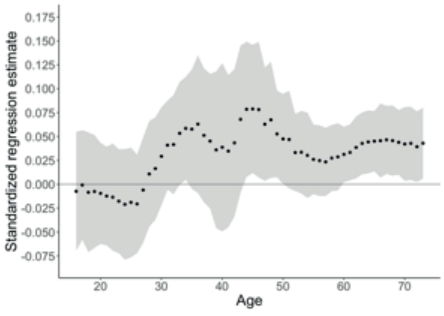


Figure S4.3. Australia: standardized regression estimates with 95% confidence intervals (as grey banners) from age weighted mixed effects models with Buss Perry subscales as outcome variables (Physical Aggression, Hostility, Anger, Verbal Aggression)

Table S4.1. The Netherlands: Results from age-weighted models

			Bootstrapped 95% CI		
AGE	**β**	**Bootstrapped β**	**Lower**	**Upper**	***N***
12	0.05	0.05	0.03	0.07	6032
13	0.05	0.05	0.02	0.08	7142
14	0.05	0.05	0.02	0.07	8324
15	0.05	0.05	0.02	0.07	9478
16	0.04	0.04	0.02	0.07	10428
17	0.04	0.04	0.02	0.06	11147
18	0.04	0.04	0.02	0.07	11506
19	0.04	0.04	0.01	0.06	11531
20	0.04	0.04	0.02	0.06	11236
21	0.04	0.04	0.02	0.06	10706
22	0.05	0.04	0.02	0.07	9997
23	0.05	0.05	0.02	0.07	9176
24	0.05	0.05	0.02	0.07	8302
25	0.05	0.05	0.03	0.07	7415
26	0.05	0.05	0.02	0.08	6559
27	0.05	0.05	0.03	0.08	5770
28	0.05	0.05	0.03	0.08	5058
29	0.05	0.05	0.02	0.07	4461
30	0.05	0.05	0.02	0.08	3980
31	0.04	0.04	0.01	0.07	3604
32	0.04	0.04	0.01	0.07	3332
33	0.04	0.04	0.01	0.07	3146
34	0.04	0.04	0.01	0.07	3005
35	0.04	0.04	0.01	0.07	2895
36	0.04	0.04	0.01	0.08	2801
37	0.04	0.05	0.01	0.08	2697
38	0.04	0.05	0.01	0.09	2570
39	0.04	0.04	0.01	0.08	2426
40	0.04	0.04	0.01	0.08	2269
41	0.03	0.04	0	0.07	2106
42	0.03	0.03	0	0.07	2083
43	0.03	0.03	-0.01	0.07	2066
44	0.02	0.02	-0.01	0.06	2061
45	0.02	0.02	-0.02	0.06	2081
46	0.02	0.02	-0.02	0.05	2129
47	0.02	0.02	-0.01	0.05	2207

48	0.01	0.01	-0.02	0.04	2309
49	0.01	0.01	-0.02	0.05	2428
50	0.01	0.01	-0.02	0.05	2553
51	0.01	0.01	-0.02	0.04	2674
52	0.01	0.01	-0.02	0.05	2783
53	0.02	0.01	-0.01	0.04	2875
54	0.02	0.02	-0.01	0.05	2943
55	0.02	0.02	-0.01	0.05	2987
56	0.02	0.02	-0.01	0.05	3006
57	0.02	0.02	-0.01	0.05	3001
58	0.02	0.02	-0.01	0.05	2968
59	0.03	0.02	-0.01	0.06	2905
60	0.03	0.03	0	0.06	2807
61	0.03	0.03	0	0.06	2676
62	0.03	0.03	-0.01	0.06	2513
63	0.02	0.02	-0.01	0.06	2317
64	0.02	0.02	-0.01	0.05	2101
65	0.02	0.02	-0.02	0.05	1876
66	0.01	0.02	-0.02	0.05	1648
67	0.01	0.01	-0.03	0.05	1429
68	0	0.01	-0.03	0.04	1224
69	0	0	-0.03	0.03	1036
70	0	0	-0.04	0.04	866

Note: β= standardized regression estimate, Bootstrapped β= Bootstrapped standardized regression estimate, Bootstrapped 95% CI= empirical 95% confidence intervals, *N*= sum of regression weights.

Table S4.2. Australia: Results from age-weighted models

AGE	β	Bootstrapped β	Bootstrapped 95% CI		*N*
			Lower	Upper	
16	0.03	0.03	-0.03	0.1	754
17	0.03	0.03	-0.03	0.09	829
18	0.03	0.02	-0.03	0.08	898
19	0.03	0.02	-0.03	0.08	968
20	0.03	0.03	-0.03	0.08	1047
21	0.02	0.02	-0.04	0.08	1140
22	0.02	0.02	-0.04	0.08	1249
23	0.02	0.01	-0.04	0.06	1373
24	0.01	0.01	-0.05	0.06	1498
25	0.01	0.01	-0.05	0.06	1611
26	0.01	0.01	-0.05	0.06	1704
27	0.01	0.01	-0.05	0.06	1770
28	0.02	0.02	-0.04	0.07	1806
29	0.03	0.03	-0.03	0.08	1808
30	0.03	0.04	-0.01	0.09	1785
31	0.04	0.04	-0.01	0.09	1744
32	0.05	0.05	-0.01	0.12	1688
33	0.06	0.06	0	0.13	1620
34	0.06	0.07	0	0.13	1526
35	0.07	0.07	-0.01	0.16	1405
36	0.07	0.08	0	0.15	1250
37	0.08	0.07	-0.01	0.15	1067
38	0.08	0.08	0	0.16	868
39	0.08	0.07	0.01	0.14	672
40	0.09	0.08	0.01	0.15	496
41	0.1	0.09	0.01	0.17	357
42	0.1	0.1	0.03	0.16	262
43	0.11	0.11	0.05	0.17	209
44	0.11	0.1	0.03	0.17	193
45	0.1	0.09	0.03	0.15	215
46	0.08	0.08	0.02	0.14	265
47	0.07	0.07	0.01	0.12	332
48	0.05	0.05	0	0.11	410
49	0.04	0.04	-0.01	0.1	498
50	0.03	0.04	-0.01	0.09	596
51	0.03	0.03	-0.02	0.08	699

52	0.02	0.03	-0.02	0.07	801
53	0.02	0.02	-0.03	0.07	900
54	0.02	0.02	-0.03	0.07	999
55	0.02	0.02	-0.03	0.07	1091
56	0.02	0.02	-0.02	0.06	1174
57	0.02	0.02	-0.02	0.06	1242
58	0.02	0.01	-0.02	0.05	1300
59	0.02	0.02	-0.02	0.06	1347
60	0.02	0.02	-0.02	0.05	1381
61	0.02	0.02	-0.02	0.06	1401
62	0.03	0.03	-0.01	0.06	1404
63	0.03	0.03	-0.01	0.07	1391
64	0.04	0.04	0	0.08	1359
65	0.04	0.04	0	0.08	1309
66	0.04	0.04	0	0.08	1245
67	0.04	0.05	0.01	0.08	1169
68	0.04	0.04	0.01	0.08	1084
69	0.04	0.04	0	0.09	993
70	0.05	0.04	0	0.09	897
71	0.05	0.04	0	0.08	797
72	0.05	0.05	0.01	0.09	695
73	0.05	0.04	-0.01	0.09	594
16	0.03	0.03	-0.03	0.1	754

Note: β= standardized regression estimate, Bootstrapped β= Bootstrapped standardized regression estimate, Bootstrapped 95% CI= empirical 95% confidence intervals, *N*= sum of regression weights.

Table S6.1. Between-/within-family sample overlap matrix

	TRF5	TRF7	TRF10	TRF12	CBCL V3	DCB V5	CBCL V7	CBCL V10	CBCL V12	CBCL M3	DCB M5	CBCL M7	CBCL M10	CBCL M12	YSR 1991	YSR 1995	YSR 1997	YSR 2000	ASR 2009	ASR 2014	YSR 14	YSR 16	YSR 18	YSR Pilot
TRF 5	140	4	74	54	36	132	122	100	100	128	132	130	120	118	0	0	0	0	31	28	89	51	0	0
TRF 7	4	1384	677	330	784	841	660	645	385	889	889	803	809	479	0	0	0	0	0	152	357	267	0	0
TRF 10	74	677	1492	462	758	1015	845	924	600	1039	1051	977	1111	722	0	0	0	0	119	249	546	423	0	0
TRF 12	54	330	462	1091	660	860	774	757	704	890	888	888	898	786	0	0	0	0	238	274	388	390	182	80
CBCL V3	36	784	758	660	1890	1662	1340	1351	945	1847	1702	1534	1603	1100	0	0	0	0	207	361	591	546	238	134
DCB V5	132	841	1015	860	1662	2384	1757	1722	1270	2128	2374	2017	2044	1488	0	0	0	0	374	501	817	761	248	148
CBCL V7	122	660	845	774	1340	1757	1937	1530	1213	1727	1781	1901	1710	1350	0	0	0	0	359	464	756	710	230	136
CBCL V10	100	645	924	757	1351	1722	1530	1956	1178	1711	1750	1640	1914	1274	0	0	0	0	366	463	641	652	259	126
CBCL V12	100	385	600	704	945	1270	1213	1178	1416	1258	1290	1293	1276	1396	0	0	0	0	362	459	690	655	250	130
CBCL M3	128	889	1039	890	1847	2128	1727	1711	1258	2524	2244	2057	2109	1507	0	0	0	0	352	495	823	775	264	148
DCB M5	132	889	1051	888	1702	2374	1781	1750	1290	2244	2524	2105	2144	1548	0	0	0	0	382	524	846	797	256	154
CBCL M7	130	803	977	888	1534	2017	1901	1640	1293	2057	2105	2343	2032	1560	0	0	0	0	381	522	864	799	260	150
CBCL M10	120	809	1111	898	1603	2044	1710	1914	1276	2109	2144	2032	2461	1516	0	0	0	0	392	519	774	773	282	146
CBCL M12	118	479	722	786	1100	1488	1350	1274	1396	1507	1548	1560	1516	1710	0	0	0	0	387	515	826	766	272	144
YSR 1991	0	0	0	0	0	0	0	0	0	0	0	0	0	0	583	420	298	284	201	180	0	0	0	0
YSR 1995	0	0	0	0	0	0	0	0	0	0	0	0	0	0	420	680	422	406	272	232	0	0	0	0
YSR 1997	0	0	0	0	0	0	0	0	0	0	0	0	0	0	298	422	1788	1197	771	540	0	0	2	0
YSR 2000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	284	406	1197	1921	969	707	0	0	2	0
ASR 2009	31	0	119	238	207	374	359	366	362	352	382	381	392	387	201	272	771	969	2193	1142	183	314	210	84
ASR 2014	28	152	249	274	361	501	464	463	459	495	524	522	519	515	180	232	540	707	1142	1985	448	555	168	74
YSR 14	89	357	546	388	591	817	756	641	690	823	846	864	774	826	0	0	0	0	183	448	1345	803	0	0
YSR 16	51	267	423	390	546	761	710	652	655	775	797	799	773	766	0	0	0	0	314	555	803	1357	113	81
YSR 18	0	0	0	182	238	248	230	259	250	264	256	260	282	272	0	0	2	2	210	168	0	113	464	143
YSR Pilot	0	0	0	80	134	148	136	126	130	148	154	150	146	144	0	0	0	0	84	74	0	81	143	228

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = ASEBA Youth Self-Report, ASR = ASEBA Adult Self-report.

Table S6.2. Between-/within-family cross-sample correlation matrix

	TRF5	TRF7	TRF10	TRF12	CBCL V3	DCB V5	CBCL V7	CBCL V10	CBCL V12	CBCL M3	DCB M5	CBCL M7	CBCL M10	CBCL M12	YSR 1991	YSR 1995	YSR 1997	YSR 2000	ASR 2009	ASR 2014	YSR 14	YSR 16	YSR 18	YSR Pilot
TRF 5	1	NA	0.11	0.18	0.27	0.12	0.14	0.25	0.29	0.11	0.16	0.18	0.24	0.19	NA	NA	NA	NA	0.23	0.32	0.26	0.1	NA	NA
TRF 7	NA	1	0.54	0.42	0.18	0.26	0.24	0.32	0.22	0.23	0.3	0.31	0.32	0.25	NA	NA	NA	NA	NA	-0.13	0.12	0.13	NA	NA
TRF 10	0.11	0.54	1	0.51	0.18	0.22	0.29	0.39	0.34	0.18	0.28	0.36	0.37	0.35	NA	NA	NA	NA	0.06	0.05	0.19	0.09	NA	NA
TRF 12	0.18	0.42	0.51	1	0.09	0.24	0.28	0.29	0.29	0.13	0.19	0.29	0.31	0.35	NA	NA	NA	NA	0.19	-0.01	0.25	0.16	0.18	0.21
CBCL V3	0.27	0.18	0.18	0.09	1	0.38	0.54	0.48	0.42	0.66	0.35	0.47	0.41	0.35	NA	NA	NA	NA	0.07	0.08	0.18	0.07	0.18	0.27
DCB V5	0.12	0.26	0.22	0.24	0.38	1	0.42	0.38	0.35	0.34	0.57	0.39	0.38	0.32	NA	NA	NA	NA	0.12	0.09	0.2	0.1	0.23	0.16
CBCL V7	0.14	0.24	0.29	0.28	0.54	0.42	1	0.66	0.57	0.43	0.4	0.69	0.56	0.5	NA	NA	NA	NA	0.21	0.11	0.27	0.18	0.21	0.17
CBCL V10	0.25	0.32	0.39	0.29	0.48	0.38	0.66	1	0.67	0.39	0.36	0.58	0.7	0.58	NA	NA	NA	NA	0.2	0.16	0.3	0.22	0.28	0.28
CBCL V12	0.29	0.22	0.34	0.29	0.42	0.35	0.57	0.67	1	0.33	0.27	0.51	0.56	0.68	NA	NA	NA	NA	0.17	0.14	0.35	0.24	0.28	0.29
CBCL M3	0.11	0.23	0.18	0.13	0.66	0.34	0.43	0.39	0.33	1	0.46	0.56	0.5	0.43	NA	NA	NA	NA	0.08	0.09	0.23	0.13	0.12	0.24
DCB M5	0.16	0.3	0.28	0.19	0.35	0.57	0.4	0.36	0.27	0.46	1	0.51	0.48	0.37	NA	NA	NA	NA	0.13	0.06	0.27	0.14	0.16	0.19
CBCL M7	0.18	0.31	0.36	0.29	0.47	0.39	0.69	0.58	0.51	0.56	0.51	1	0.71	0.65	NA	NA	NA	NA	0.19	0.16	0.32	0.21	0.24	0.21
CBCL M10	0.24	0.32	0.37	0.31	0.41	0.38	0.56	0.7	0.56	0.5	0.48	0.71	1	0.7	NA	NA	NA	NA	0.18	0.22	0.36	0.22	0.27	0.26
CBCL M12	0.19	0.25	0.35	0.35	0.35	0.32	0.5	0.58	0.68	0.43	0.37	0.65	0.7	1	NA	NA	NA	NA	0.21	0.3	0.4	0.27	0.32	0.3
YSR 1991	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	0.46	0.38	0.32	0.27	0.19	NA	NA	NA	NA
YSR 1995	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.46	1	0.57	0.48	0.36	0.35	NA	NA	NA	NA
YSR 1997	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.38	0.57	1	0.62	0.49	0.51	NA	NA	1	NA
YSR 2000	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.32	0.48	0.62	1	0.51	0.51	NA	NA	1	NA
ASR 2009	0.23	NA	0.06	0.19	0.07	0.12	0.21	0.2	0.17	0.08	0.13	0.19	0.18	0.21	0.27	0.36	0.49	0.51	1	0.62	0.41	0.48	0.49	0.36
ASR 2014	0.32	-0.13	0.05	-0.01	0.08	0.09	0.11	0.16	0.14	0.09	0.06	0.16	0.22	0.3	0.19	0.35	0.51	0.51	0.62	1	0.32	0.38	0.33	0.22
YSR 14	0.26	0.12	0.19	0.25	0.18	0.2	0.27	0.3	0.35	0.23	0.27	0.32	0.36	0.4	NA	NA	NA	NA	0.41	0.32	1	0.55	NA	NA
YSR 16	0.1	0.13	0.09	0.16	0.07	0.1	0.18	0.22	0.24	0.13	0.14	0.21	0.22	0.27	NA	NA	NA	NA	0.48	0.38	0.55	1	0.67	0.71
YSR 18	NA	NA	NA	0.18	0.18	0.23	0.21	0.28	0.28	0.12	0.16	0.24	0.27	0.32	NA	NA	1	1	0.49	0.33	NA	0.67	1	0.56
YSR Pilot	NA	NA	NA	0.21	0.27	0.16	0.17	0.28	0.29	0.24	0.19	0.21	0.26	0.3	NA	NA	NA	NA	0.36	0.22	NA	0.71	0.56	1

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = ASEBA Youth Self-Report, ASR = ASEBA Adult Self-report.

Table S6.3. Transmitted/non-transmitted sample overlap matrix

	TRF5	TRF7	TRF10	TRF12	CBCL V3	DCB V5	CBCL V7	CBCL V10	CBCL V12	CBCL M3	DCB M5	CBCL M7	CBCL M10	CBCL M12	YSR 1991	YSR 1995	YSR 1997	YSR 2000	ASR 2009	ASR 2014	YSR 14	YSR 16	YSR 18	YSR Pilot
TRF 5	122	9	84	58	45	107	113	106	99	116	106	115	117	104	0	0	0	0	60	56	94	75	0	0
TRF 7	9	1687	912	494	1175	1243	829	877	389	1327	1299	1021	1072	481	0	0	0	0	0	180	359	276	0	0
TRF 10	84	912	1886	670	1165	1489	1120	1257	684	1532	1529	1287	1488	811	0	0	0	0	222	396	623	544	0	0
TRF 12	58	494	670	1398	897	1178	1034	1018	830	1203	1209	1196	1183	919	0	0	0	0	365	454	477	504	190	81
CBCL V3	45	1175	1165	897	2556	2254	1598	1687	923	2502	2283	1815	1966	1058	0	0	0	0	285	504	621	570	240	103
DCB V5	107	1243	1489	1178	2254	3238	2107	2182	1377	2851	3229	2432	2546	1579	0	0	0	0	598	798	901	935	293	117
CBCL V7	113	829	1120	1034	1598	2107	2343	1863	1281	2059	2131	2309	2075	1410	0	0	0	0	587	760	835	869	272	115
CBCL V10	106	877	1257	1018	1687	2182	1863	2499	1246	2169	2222	2026	2455	1348	0	0	0	0	584	752	727	821	299	115
CBCL V12	99	389	684	830	923	1377	1281	1246	1519	1339	1397	1383	1339	1502	0	0	0	0	562	713	779	806	279	116
CBCL M3	116	1327	1532	1203	2502	2851	2059	2169	1339	3347	2949	2439	2586	1542	0	0	0	0	565	767	897	916	281	109
DCB M5	106	1299	1529	1209	2283	3229	2131	2222	1397	2949	3353	2499	2622	1618	0	0	0	0	605	818	922	968	295	119
CBCL M7	115	1021	1287	1196	1815	2432	2309	2026	1383	2439	2499	2788	2429	1600	0	0	0	0	623	837	935	962	299	119
CBCL M10	117	1072	1488	1183	1966	2546	2075	2455	1339	2586	2622	2429	3007	1535	0	0	0	0	623	831	842	930	314	121
CBCL M12	104	481	811	919	1058	1579	1410	1348	1502	1542	1618	1600	1535	1770	0	0	0	0	607	797	890	926	302	119
YSR 1991	0	0	0	0	0	0	0	0	0	0	0	0	0	0	755	596	454	463	403	331	0	0	0	0
YSR 1995	0	0	0	0	0	0	0	0	0	0	0	0	0	0	596	973	679	671	540	432	0	0	0	0
YSR 1997	0	0	0	0	0	0	0	0	0	0	0	0	0	0	454	679	1269	941	713	582	0	0	2	1
YSR 2000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	463	671	941	1329	807	678	0	0	2	1
ASR 2009	60	0	222	365	285	598	587	584	562	565	605	623	623	607	403	540	713	807	1967	1329	265	440	251	88
ASR 2014	56	180	396	454	504	798	760	752	713	767	818	837	831	797	331	432	582	678	1329	1978	532	678	221	81
YSR 14	94	359	623	477	621	901	835	727	779	897	922	935	842	890	0	0	0	0	265	532	1117	725	0	0
YSR 16	75	276	544	504	570	935	869	821	806	916	968	962	930	926	0	0	0	0	440	678	725	1186	109	46
YSR 18	0	0	0	190	240	293	272	299	279	281	295	299	314	302	0	0	2	2	251	221	0	109	415	132
YSR Pilot	122	9	84	58	45	107	113	106	99	116	106	115	117	104	0	0	0	0	60	56	94	75	0	0

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = ASEBA Youth Self-Report, ASR = ASEBA Adult Self-report.

Table S6.4. Transmitted/non-transmitted cross-sample correlation matrix

	TRF5	TRF7	TRF10	TRF12	CBCL V3	DCB V5	CBCL V7	CBCL V10	CBCL V12	CBCL M3	DCB M5	CBCL M7	CBCL M10	CBCL M12	YSR 1991	YSR 1995	YSR 1997	YSR 2000	ASR 2009	ASR 2014	YSR 14	YSR 16	YSR 18	YSR Pilot
TRF 5	1	0.66	0.22	0.11	0.35	0.15	0.15	0.07	0.09	0.01	0.21	0.2	0.1	0.17	NA	NA	NA	NA	0.15	-0.21	0.02	-0.07	NA	NA
TRF 7	0.66	1	0.49	0.36	0.16	0.3	0.27	0.28	0.25	0.25	0.31	0.33	0.31	0.34	NA	NA	NA	NA	NA	-0.02	0.11	0.03	NA	NA
TRF 10	0.22	0.49	1	0.49	0.14	0.23	0.29	0.32	0.26	0.18	0.29	0.3	0.31	0.31	NA	NA	NA	NA	0.06	0.1	0.14	0.06	NA	NA
TRF 12	0.11	0.36	0.49	1	0.08	0.25	0.31	0.28	0.22	0.15	0.22	0.28	0.32	0.32	NA	NA	NA	NA	0.11	0.06	0.16	0.14	0.15	0.2
CBCL V3	0.35	0.16	0.14	0.08	1	0.35	0.51	0.45	0.44	0.68	0.32	0.46	0.4	0.36	NA	NA	NA	NA	0.2	0.08	0.2	0.09	0.18	0.26
DCB V5	0.15	0.3	0.23	0.25	0.35	1	0.44	0.42	0.39	0.35	0.6	0.41	0.4	0.33	NA	NA	NA	NA	0.17	0.09	0.24	0.19	0.14	0.22
CBCL V7	0.15	0.27	0.29	0.31	0.51	0.44	1	0.67	0.63	0.46	0.41	0.71	0.59	0.55	NA	NA	NA	NA	0.27	0.15	0.34	0.24	0.14	0.01
CBCL V10	0.07	0.28	0.32	0.28	0.45	0.42	0.67	1	0.7	0.41	0.39	0.59	0.71	0.62	NA	NA	NA	NA	0.21	0.18	0.34	0.3	0.3	0.17
CBCL V12	0.09	0.25	0.26	0.22	0.44	0.39	0.63	0.7	1	0.4	0.33	0.55	0.57	0.7	NA	NA	NA	NA	0.2	0.17	0.36	0.32	0.28	0.12
CBCL M3	0.01	0.25	0.18	0.15	0.68	0.35	0.46	0.41	0.4	1	0.46	0.58	0.5	0.48	NA	NA	NA	NA	0.24	0.13	0.29	0.19	0.19	0.23
DCB M5	0.21	0.31	0.29	0.22	0.32	0.6	0.41	0.39	0.33	0.46	1	0.52	0.47	0.41	NA	NA	NA	NA	0.23	0.1	0.27	0.19	0.13	0.25
CBCL M7	0.2	0.33	0.3	0.28	0.46	0.41	0.71	0.59	0.55	0.58	0.52	1	0.71	0.65	NA	NA	NA	NA	0.24	0.12	0.39	0.26	0.23	0.16
CBCL M10	0.1	0.31	0.31	0.32	0.4	0.4	0.59	0.71	0.57	0.5	0.47	0.71	1	0.71	NA	NA	NA	NA	0.25	0.17	0.36	0.29	0.32	0.19
CBCL M12	0.17	0.34	0.31	0.32	0.36	0.33	0.55	0.62	0.7	0.48	0.41	0.65	0.71	1	NA	NA	NA	NA	0.28	0.22	0.45	0.3	0.31	0.26
YSR 1991	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	0.5	0.4	0.42	0.23	0.21	NA	NA	NA	NA
YSR 1995	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.5	1	0.61	0.55	0.36	0.37	NA	NA	NA	NA
YSR 1997	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.4	0.61	1	0.61	0.47	0.46	NA	NA	1	NA
YSR 2000	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.42	0.55	0.61	1	0.49	0.49	NA	NA	1	NA
ASR 2009	0.15	NA	0.06	0.11	0.2	0.17	0.27	0.21	0.2	0.24	0.23	0.24	0.25	0.28	0.23	0.36	0.47	0.49	1	0.59	0.42	0.49	0.47	0.35
ASR 2014	-0.21	-0.02	0.1	0.06	0.08	0.09	0.15	0.18	0.17	0.13	0.1	0.12	0.17	0.22	0.21	0.37	0.46	0.49	0.59	1	0.32	0.39	0.29	0.2
YSR 14	0.02	0.11	0.14	0.16	0.2	0.24	0.34	0.34	0.36	0.29	0.27	0.39	0.36	0.45	NA	NA	NA	NA	0.42	0.32	1	0.59	NA	NA
YSR 16	-0.07	0.03	0.06	0.14	0.09	0.19	0.24	0.3	0.32	0.19	0.19	0.26	0.29	0.3	NA	NA	NA	NA	0.49	0.39	0.59	1	0.65	0.71
YSR 18	NA	NA	NA	0.15	0.18	0.14	0.14	0.3	0.28	0.19	0.13	0.23	0.32	0.31	NA	NA	1	1	0.47	0.29	NA	0.65	1	0.57
YSR Pilot	NA	NA	NA	0.2	0.26	0.22	0.01	0.17	0.12	0.23	0.25	0.16	0.19	0.26	NA	NA	NA	NA	0.35	0.2	NA	0.71	0.57	1

Abbreviations: TRF = Teacher Report Form, CBCL = Child Behavior Checklist, DCB = Devereux Child Behavior rating scale, YSR = ASEBA Youth Self-Report, ASR = ASEBA Adult Self-report, NA= not available (no cross-sample overlap).

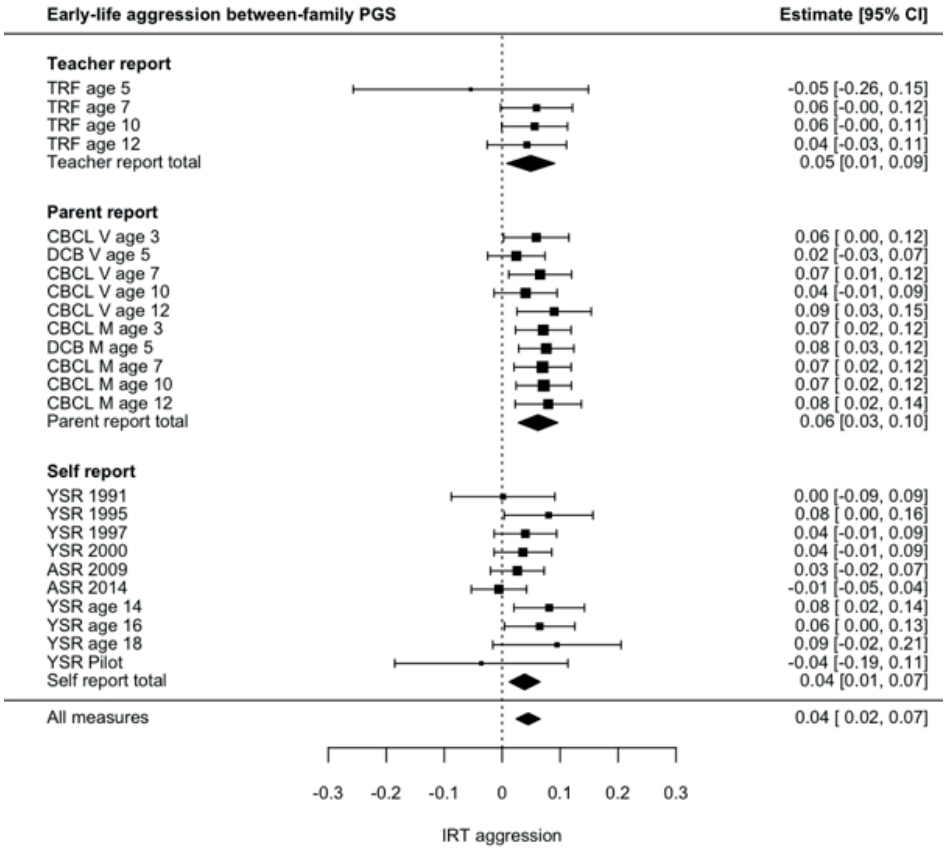


Figure S6.1. EL-AGG PGS between-family results

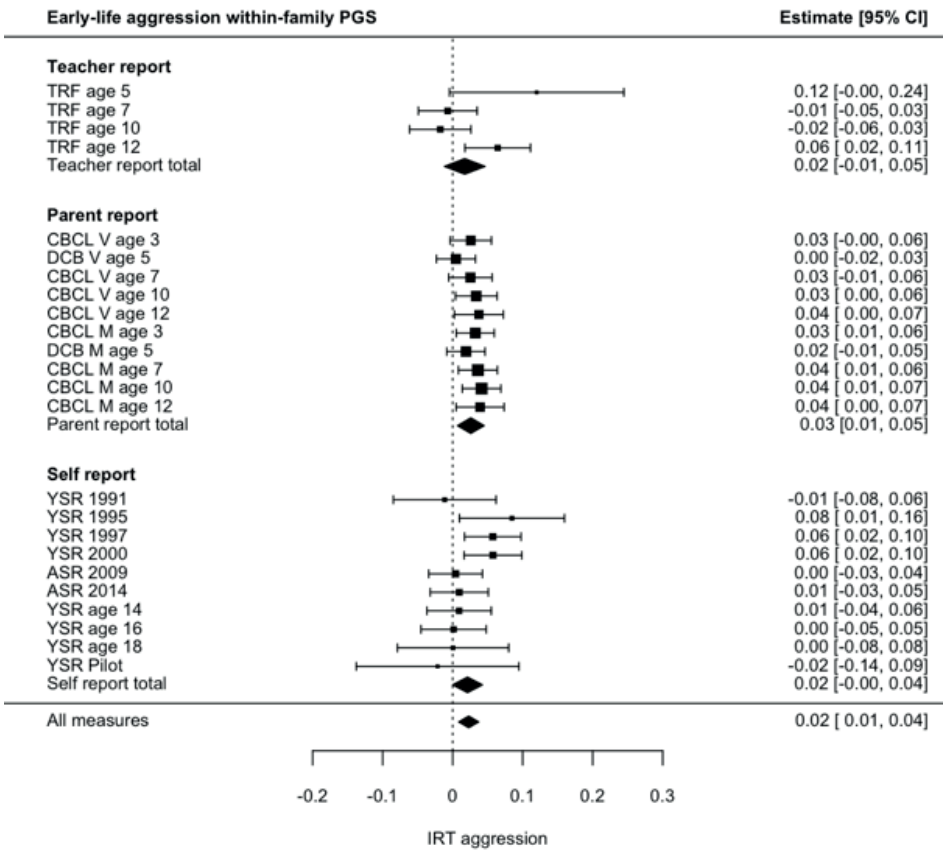


Figure S6.2. EL-AGG PGS within-family results

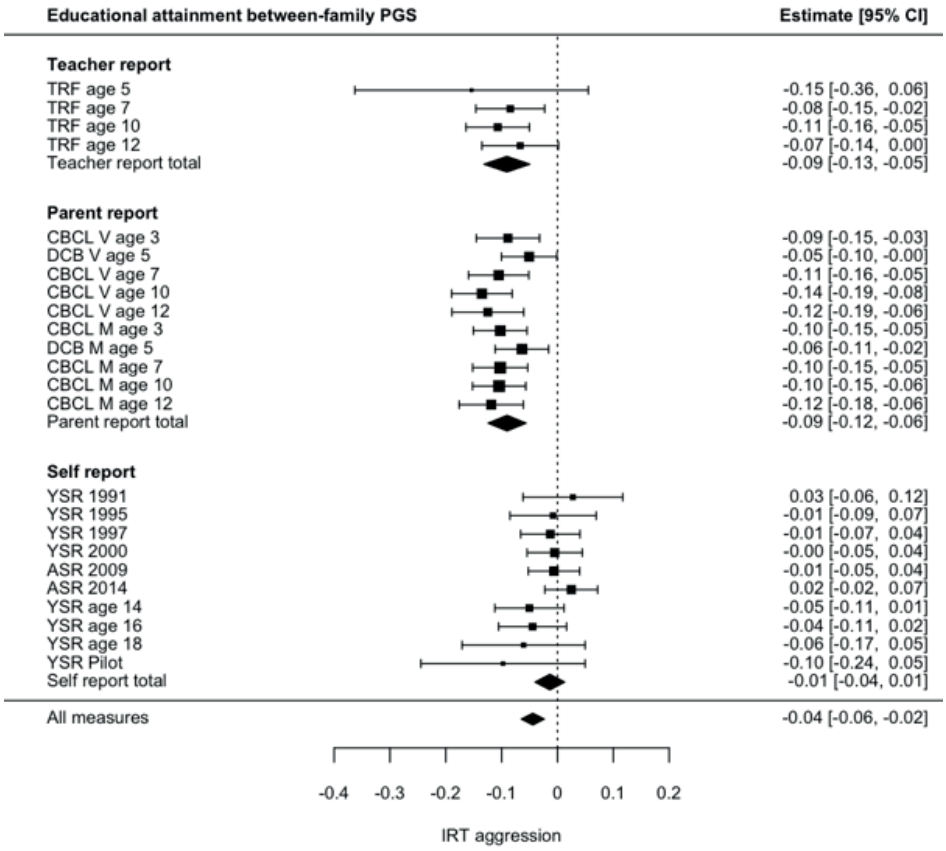


Figure S6.3. EA PGS between-family results

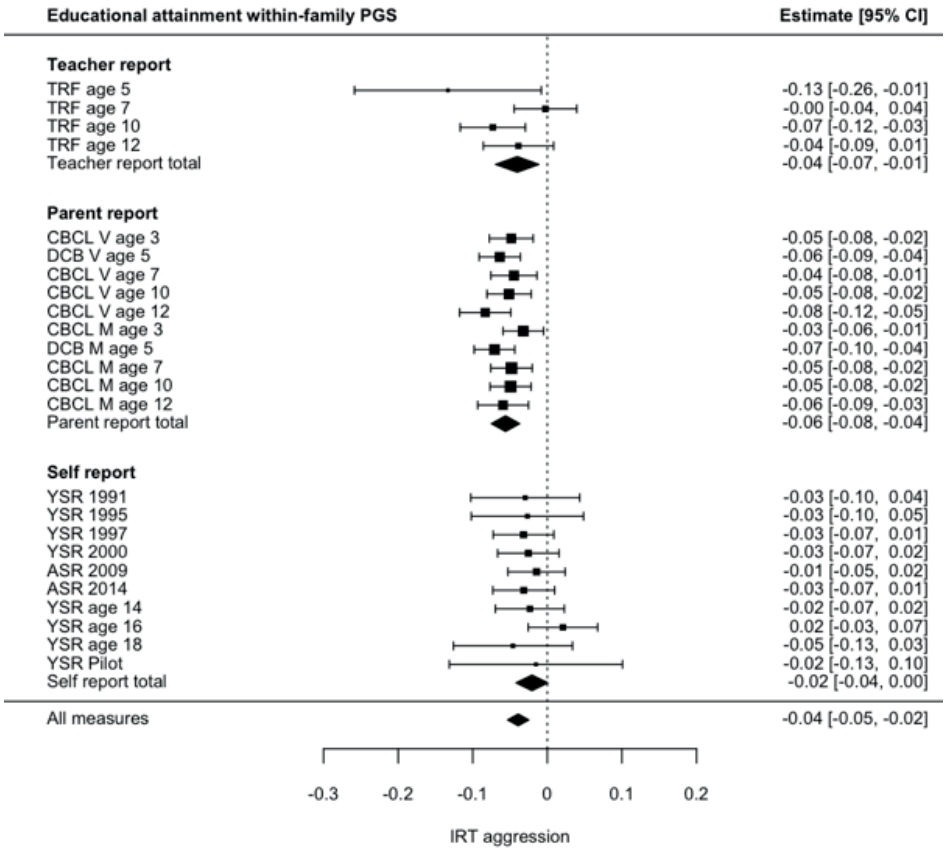


Figure S6.4. EA PGS within-family results

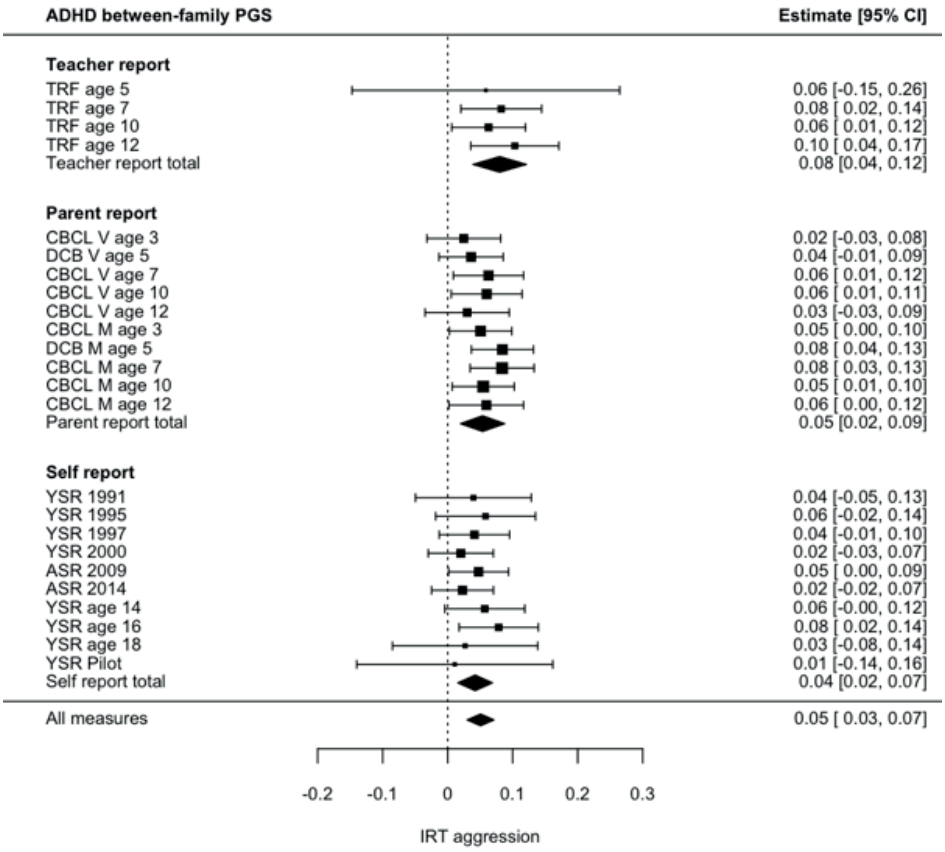


Figure S6.5. ADHD PGS between-family results

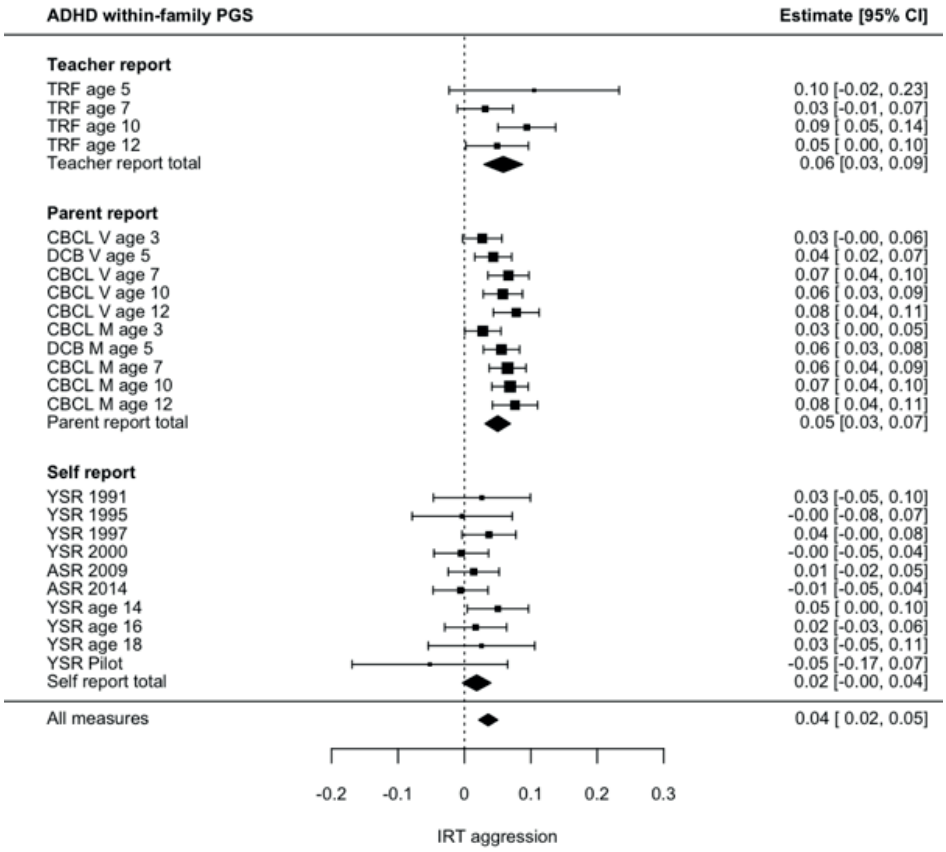


Figure S6.6. ADHD PGS within-family results

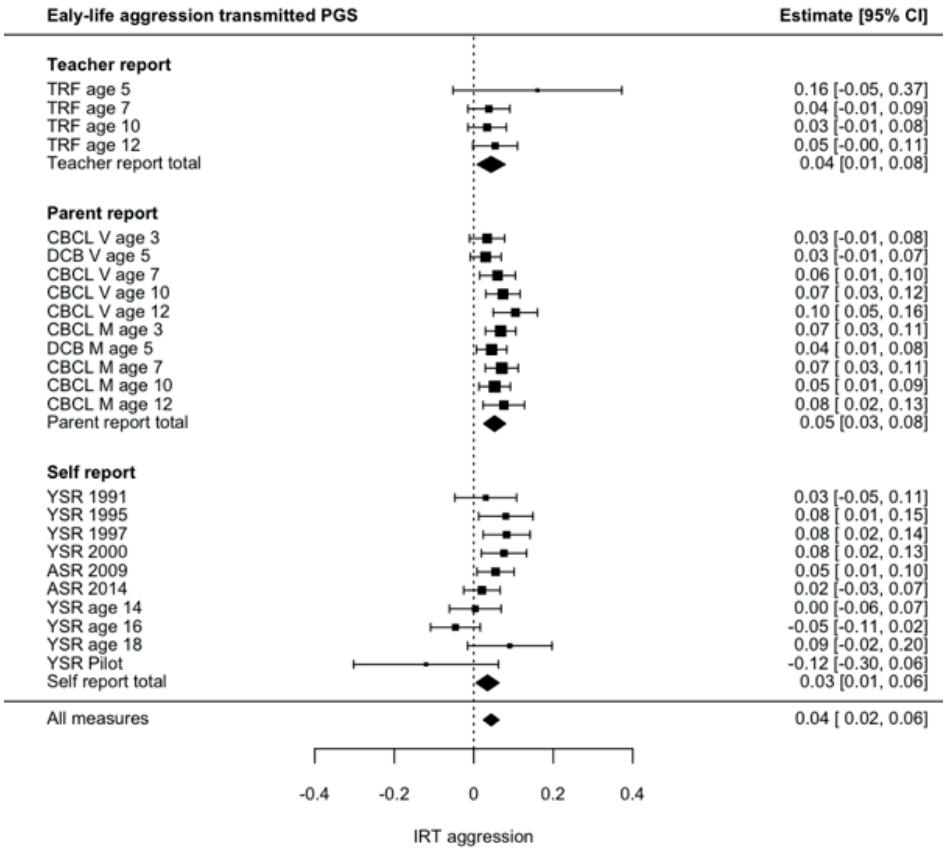


Figure S6.7. EL-AGG transmitted PGS results

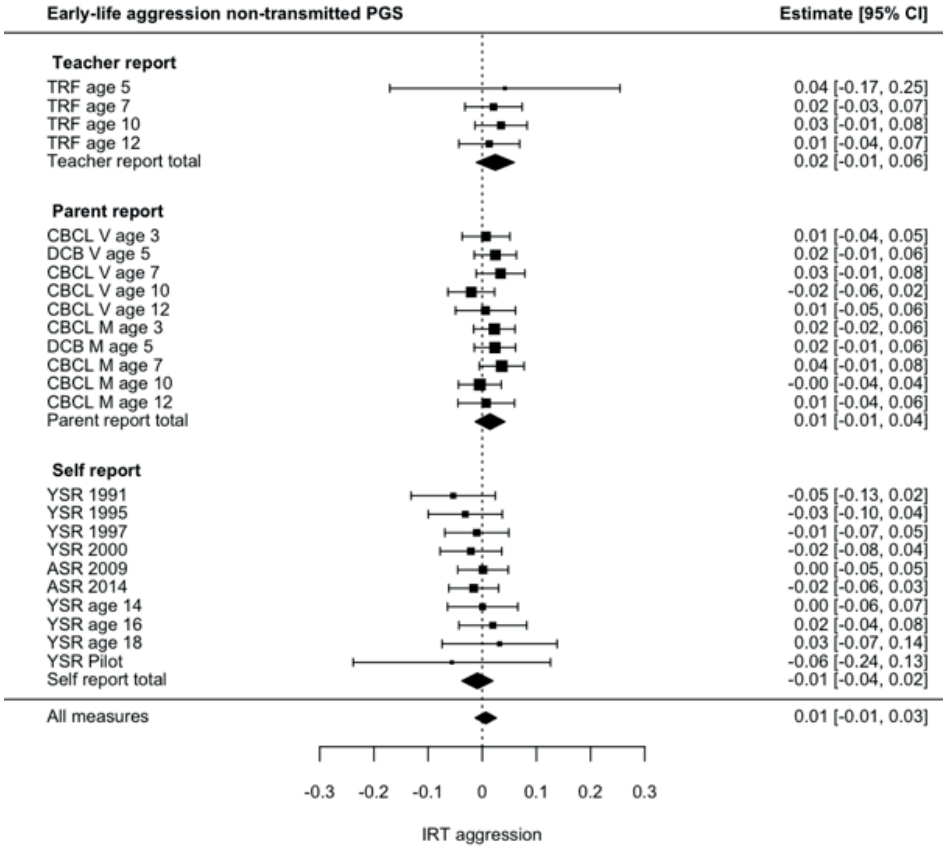


Figure S6.8. EL-AGG non-transmitted PGS results

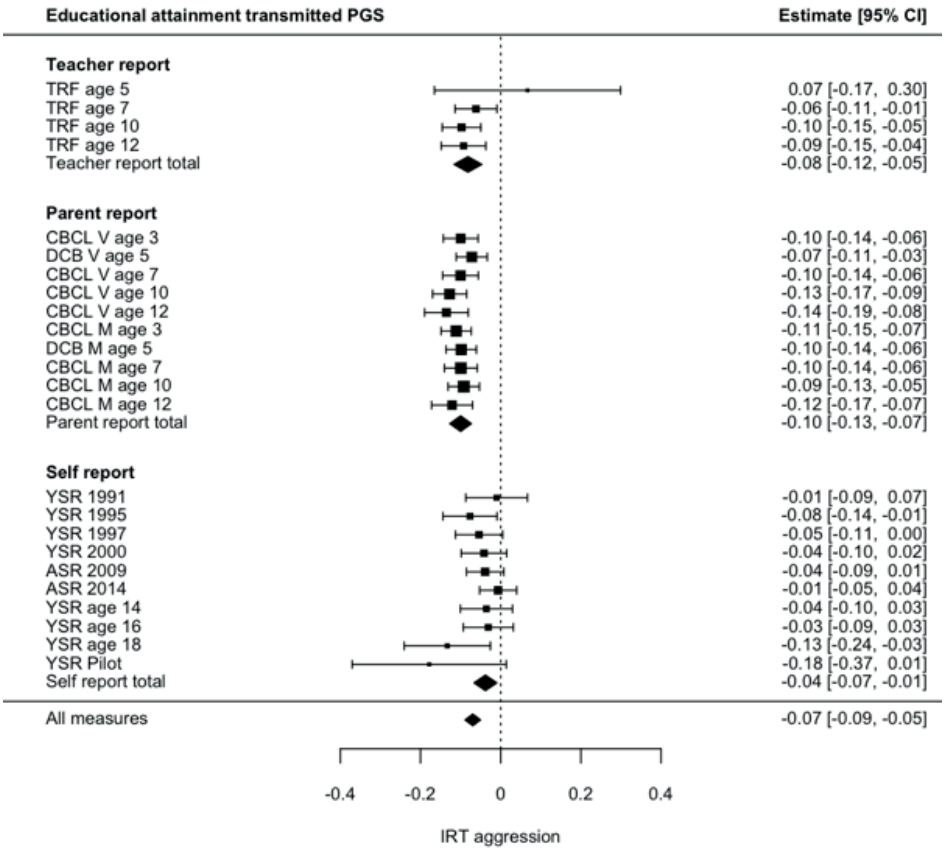


Figure S6.9. EA transmitted PGS results

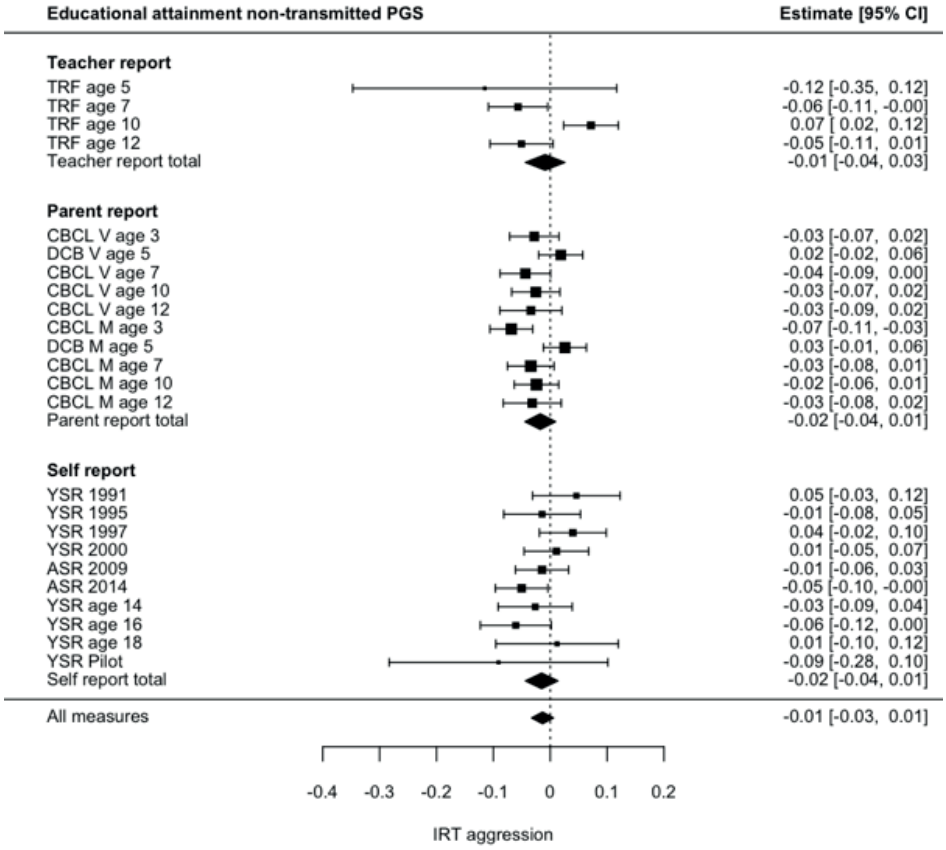


Figure S6.10. EA non-transmitted PGS results

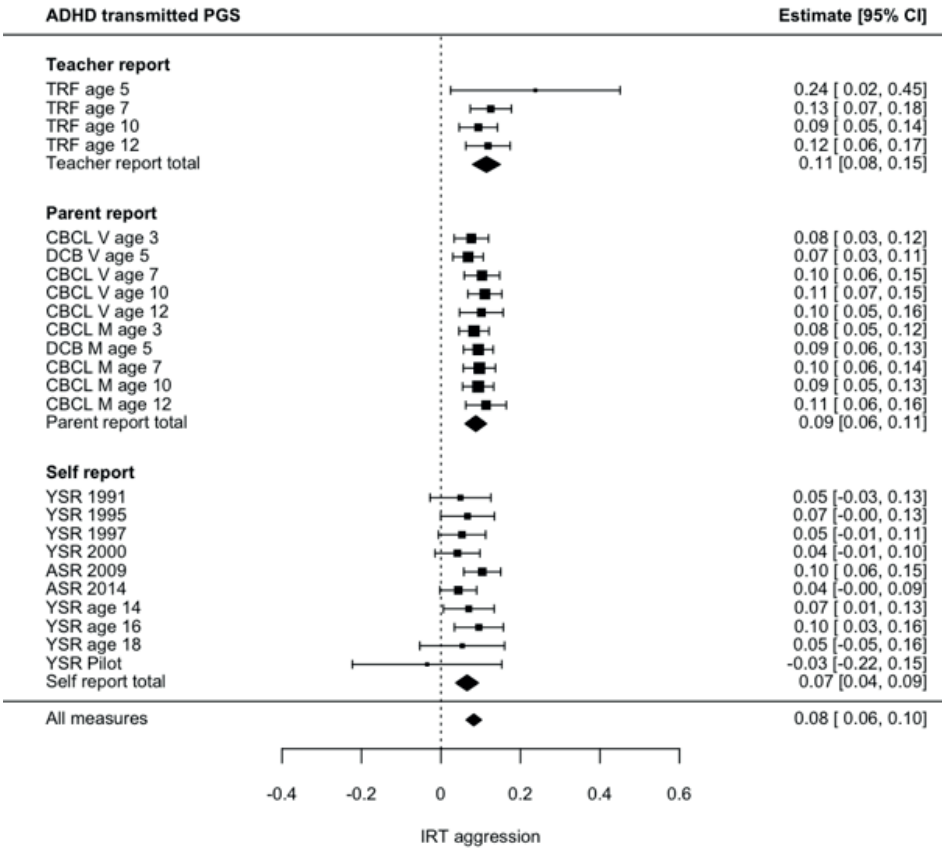


Figure S6.11. ADHD transmitted PGS results

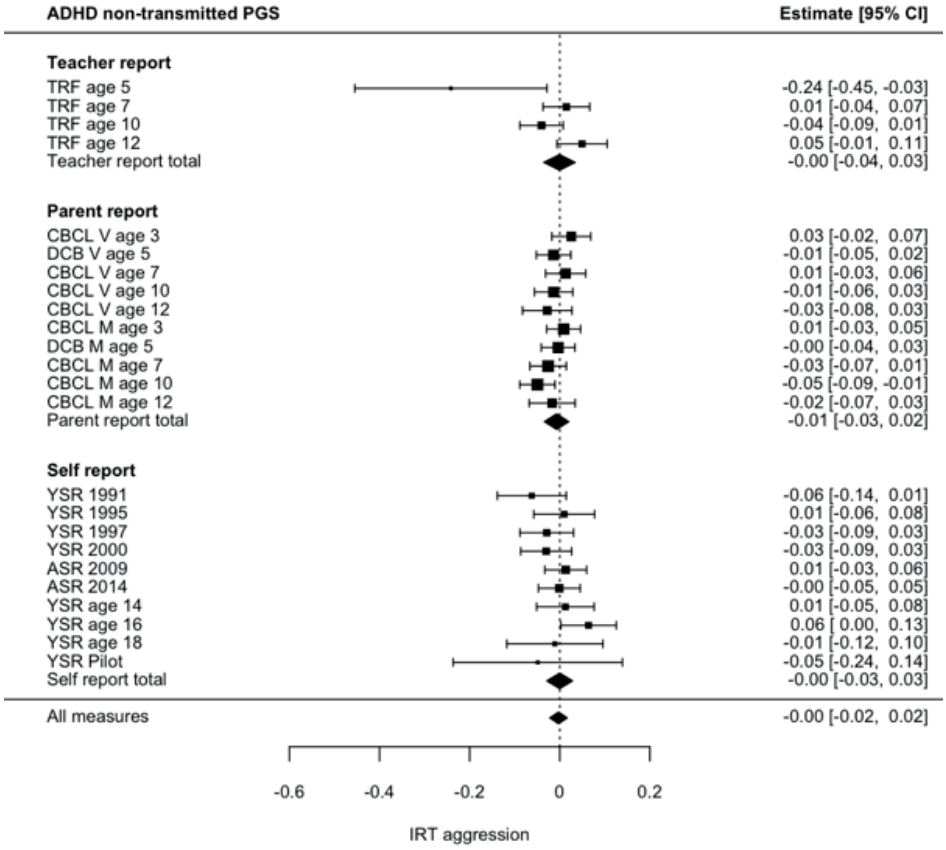
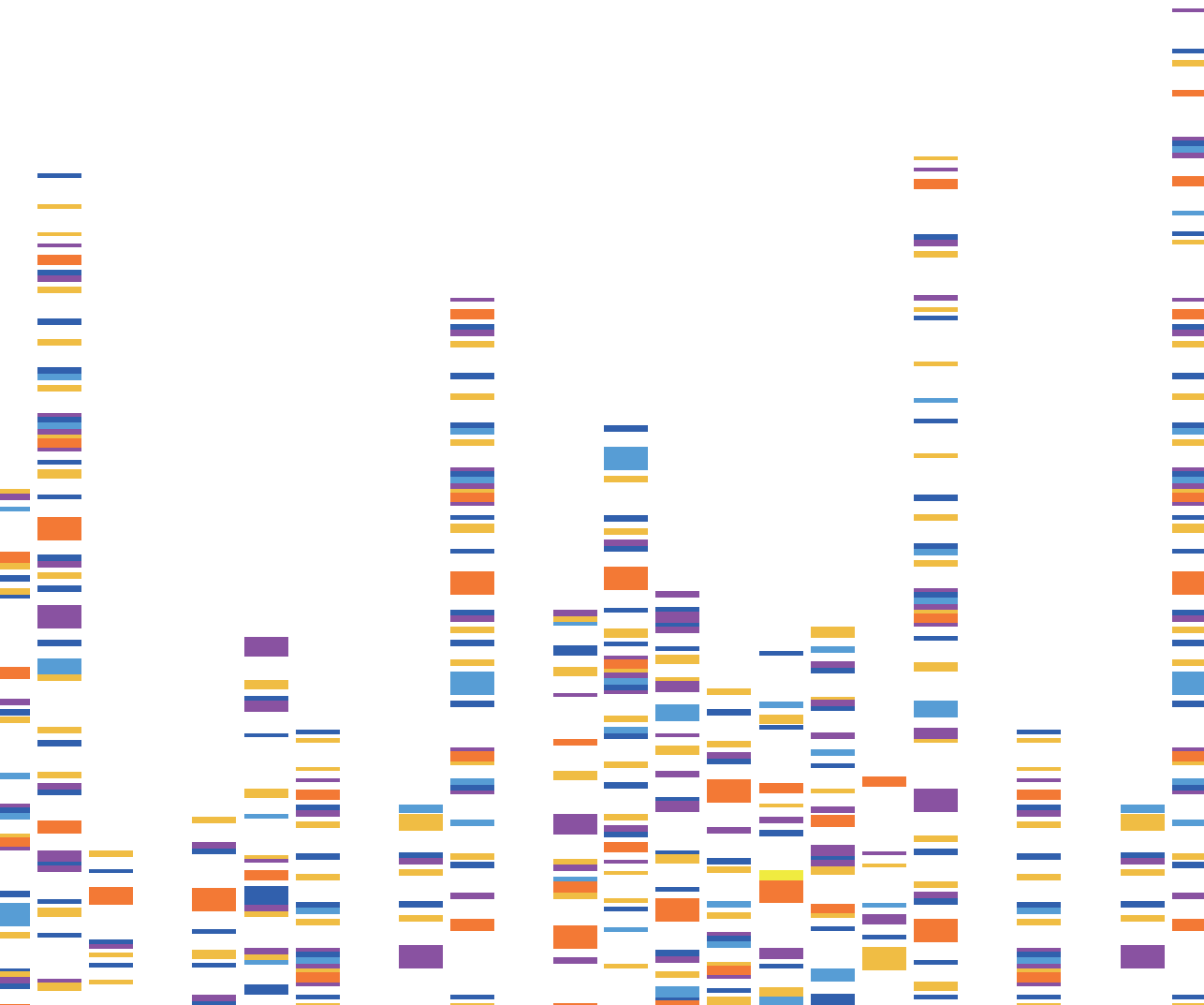
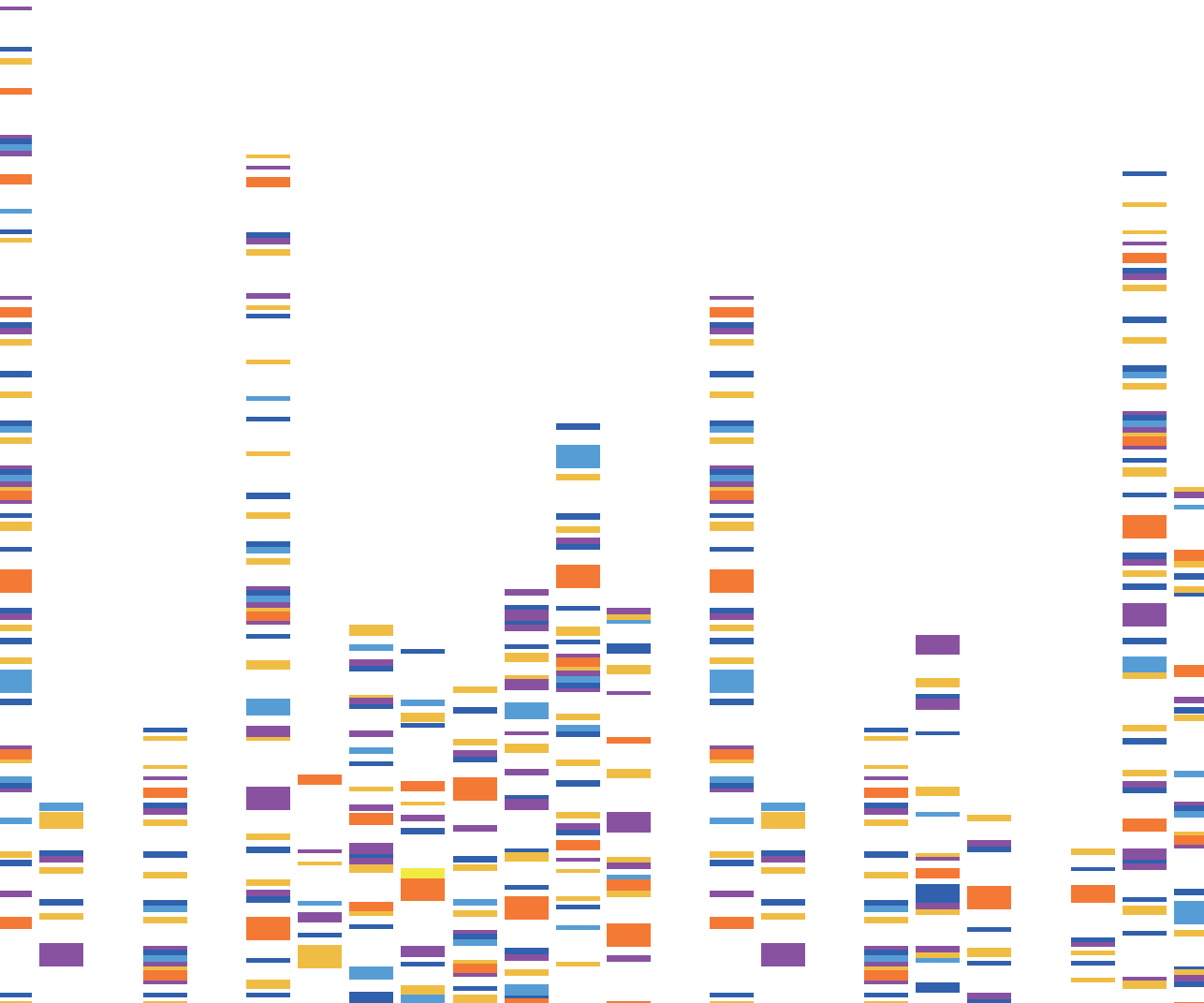


Figure S6.12. ADHD non-transmitted PGS results



11
Appendix



ABOUT THE AUTHOR

Camiel van der Laan obtained his bachelor degree in 'Algemene Sociale Wetenschappen' at the University of Amsterdam in 2013. In 2016 he completed the research master Child Development and Education at the University of Amsterdam. Camiel started his doctoral research in 2017, studying genetic and environmental effects on aggression. He has presented his work at (inter)national conferences, including at the 2018 and 2019 International Statistical Genetics Workshop in Boulder, Colorado, the 2018 and 2019 Nederlandse Vereniging voor Criminologie conference, the 2018 European Society of Criminology conference in Sarajevo, Bosnia and Herzegovina, the 2020 conference of the Behavior Genetics Association, and the 2021 International Congress on Twin Studies. Camiel has been awarded an Amsterdam Law and Behavior Institute (A-LAB) travel grant for 1500 euros in 2019. From August 2021 he works as a postdoctoral researcher on the ODISSEI project, which aims at improving the infrastructure for social science in The Netherlands.

LIST OF PUBLICATIONS

Ip, H. F., **van der Laan, C. M.**, Brikell, I., Sánchez-Mora, C., Nolte, I. M., St. Pourcain, B., Bolhuis, K., Palviainen, T., Zafarmand, H., Colodro-Conde, L., Gordon, S., Zayats, T., Aliev, F., Jiang, C., Wang, C. A., Saunders, G., Karhunen, V., Hammerschlag, A. R., ... Boomsma, D. I. (2021). Genetic association study of childhood aggression across raters, instruments, and age. *Translational Psychiatry, 11*, Article 413. https://doi.org/10.1038/s41398-021-01480-x

Odintsova, V. V., Roetman, P. J., Ip, H. F., Pool Rene˘, **van der Laan C. M.**, Tona, K. D., Vermeiren, R. R. J. M., & Boomsma, D. I. (2019). Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool. *Psychiatric Genetics, 29*(5), 170-190. https://doi.org/10.1097/YPG.0000000000000239

Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2019). Een tweelingstudie naar indicatoren van genetische en culturele transmissie. *Tijdschrift Voor Criminologie*, *61*(2), 148-164. https://doi.org/10.5553/TvC/0165182X2019061002003

Van der Laan, C. M., Morosoli-Garcia, J. J., van de Weijer, S. G. A., Colodro-Conde, L., ACTION consortium, Lupton, M. K., ... Boomsma, D. I. (2021). Continuity of genetic risk for aggressive behavior across the life-course. *Behavior Genetics, 51*, 592-606. https://doi.org/10.1007/s10519-021-10076-6

Van der Laan, C. M., van de Weijer, S. G. A, Nivard, M. G., & Boomsma, D. I. (2021). Familial Clustering of Trends in Aggression. *Journal of Quantitative Criminology. https://doi.org/10.1007/s10940-021-09523-8*

Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2021). Genetische onderzoekdesigns in de criminologie: Een 'toolbox' voor onderzoek naar causaliteit en intergenerationele transmissie. *Tijdschrift Voor Criminologie*, *63*(3), 347-358. https://doi.org/10.5553/TvC/0165182X2021063003004

(CO)AUTHOR CONTRIBUTIONS

Chapter 2
Chapter 2 was based on an article written by C.M. van der Laan in collaboration with S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma.

Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2021). Genetische onderzoekdesigns in de criminologie: Een 'toolbox' voor onderzoek naar causaliteit en intergenerationele transmissie. *Tijdschrift Voor Criminologie*, *63*(3), 347-358. https://doi.org/10.5553/TvC/0165182X2021063003004

Chapter 3
Chapter 3 was based on an article written by C.M. van der Laan in collaboration with S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma. Analyses were done by C.M. van der Laan. The study was designed by C.M. van der Laan, S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma.

Van der Laan, C. M., van de Weijer, S. G. A, Nivard, M. G., & Boomsma, D. I. (2021). Familial Clustering of Trends in Aggression. *Journal of Quantitative Criminology. https://doi.org/10.1007/s10940-021-09523-8*

Chapter 4
Chapter 4 was based on an article written by C.M. van der Laan in collaboration with S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma, except for the methods specifically concerning the Australian cohort, which were written by J. J. Morosoli-Garcia. The study was designed by C.M. van der Laan, S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma. Analyses were done by C.M. van der Laan for the Dutch cohort, and by J. J. Morosoli-Garcia for the Australian cohort. GWAS summary statistics for computing PGSs were provided by the Action consortium. For the Dutch cohort: Genotype QC was done by J. Hottenga, PGS construction was done by R. Pool. For the Australian cohort: Genotype QC and PGS construction was done by B. Mitchell, data collection was done by M.K. Lupton, K. McAloney, R. Parker, J.M. Burns, I.B. Hickie, N.G. Martin, and S.E. Medland. All co-authors had an active role in editing and reviewing the article.

Van der Laan, C. M., Morosoli-Garcia, J. J., van de Weijer, S. G. A., Colodro-Conde, L., ACTION consortium, Lupton, M. K., ... Boomsma, D. I. (2021). Continuity of genetic risk for aggressive behavior across the life-course. *Behavior Genetics, 51*, 592-606. https://doi.org/10.1007/s10519-021-10076-6

Chapter 5

Chapter 5 was based on an article written by C.M. van der Laan in collaboration with S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma. Analyses were done by C.M. van der Laan. The study was designed by C.M. van der Laan, S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma.

Van der Laan, C. M., van de Weijer, S. G. A., Nivard, M. G., & Boomsma, D. I. (2019). Een tweelingstudie naar indicatoren van genetische en culturele transmissie. *Tijdschrift Voor Criminologie*, *61*(2), 148-164. https://doi.org/10.5553/TvC/0165182X2019061002003

Chapter 6

Chapter 6 was based on an article written by C.M. van der Laan in collaboration with S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma. Analyses and meta-analyses were done by C.M. van der Laan. Genotype QC was done by J. Hottenga, PGS construction was done by R. Pool. The study was designed by C.M. van der Laan, S.G.A. van de Weijer, M.G. Nivard, and D.I. Boomsma.

Van der Laan, C.M., van de Weijer, S.G.A., Pool, R., Hottenga, J., Nivard, M.G., & Boomsma, D.I. (submitted for publication). Direct and indirect genetic effects on aggression.

DANKWOORD

Allereerst dank aan alle deelnemers van het Nederlands Tweelingen Register. Het is een voorrecht om met deze data te mogen werken. Dank aan A-LAB voor de financiering voor dit promotieproject, en voor een week Boulder workshop.

Vier jaar zijn voorbijgevlogen. Het was zo nu en dan nogal hectisch, soms qua werk/onderzoek, maar zeker ook persoonlijk. Ik wil daarom graag alle collega's bedanken voor de steun en begrip, wat het zelfs in moeilijke tijden fijn maakte om naar de VU/ het NSCR te komen.

Speciaal dank aan mijn (co)promotoren, Dorret, Steve, en Michel. Ik heb me vanaf het begin gelukkig gevoeld met jullie als begeleiders. Dorret, jij lijkt overal wel wat vanaf te weten. Ik heb uiteraard veel gehad aan jouw altijd eerlijke input en feedback, maar jouw ogenschijnlijk onbegrensde interesse in de wetenschap, en dan heb ik het zeker niet alleen over de (gedrags) genetica, werkt aanstekelijk en inspirerend. Michel, toen ik op de VU begon werd er van alle kanten aan je getrokken om nieuwe promovendi te begeleiden. Ik snap nu waarom. Jouw ontspannen houding, gecombineerd met heel veel methodologische kennis zijn voor promovendi bijna onmisbaar. Hoewel je ook belangrijk bent voor promovendi waarbij je geen copromotor bent, ben ik toch blij dat ik je niet met een pot drop hoefde te lokken. Steve, dank voor je altijd scherpe commentaar, maar ook voor de gezelligheid. Ik denk dat we helemaal aan het begin van de vier jaar een beetje in het zelfde schuitje zaten, met de summer school en de Boulder workshop, en ik ben blij dat jij de gene was die de criminologische invalshoek gedurende de vier jaar bewaarde.

Ik heb me altijd erg thuis gevoeld op de BioPsy afdeling, niet in de minste plaats door Natascha en Michiel, jullie zijn top! Maar ook dankzij Bart en KJ, die me de eerste dag al (met verbazingwekkende precisie) het reilen en wijlen van de afdeling hebben uitgelegd. Conor, dank voor de uitnodiging te helpen bij de OpenMx werkgroepen, het helpt heel erg om meteen betrokken te worden bij de werkzaamheden van de afdeling. Also thanks to all my roommates: Sofieke, Zenab, Perline, Wonu, Eshim, Hekmat, Saskia, Veronika, Denise, and to those who made two weeks of Boulder awesome: Veronika, Eveline, Matthijs, Saskia, Perline, Margot, and Floris. Matthijs en Hill, jullie waren het wel gewend, maar jullie hulp met alle falende LISA

scripts was echt superfijn, en heeft me uren aan frustratie gescheeld. En dan nog dank voor alle gezelligheid aan de Phders van de andere kamer: Yayouk, Anne, Fiona, Margot en Lianne.

I also want to thank all NSCR colleagues, for the input and fresh views on behavior genetics/biosocial criminology, but mostly for the gezelligheid and ping-pong and Blue Angel adventures. Especially also all who joined for the conference in Sarajevo, and my roommates through the years: Janique, Lily, Iris, Sjoukje (and all of those who had no real room and were condemned to the open spaces).

And thanks to my Australian colleagues, especially José, for their collaboration on Chapter 4. My paranymphs, Wonu and Sjoukje, thanks for keeping the conversations away from science and very much focused on sports. Monika en Peter, als ik Peter moet geloven gaan jullie tijdens mijn verdediging op zoek naar jullie namen, dus bij deze. Of het nou genetica of omgeving is, iets van intergenerationele overdr... continuïteit hebben we wel kunnen demonstreren. Emma, dank voor het relativeren van het belang van promotieonderzoek, maar laat het in de toekomst maar achterwege. Na bijna 10 jaar is het in ieder geval nog steeds nooit saai.