Risk factors for disruptive behaviour: Triangulating evidence from causal inference methods

Lucy Karwatowska

A thesis submitted for the degree of

Doctorate of Philosophy

Epidemiology (Population Health Data Science)

UCL Great Ormond Street Institute of Child Health

Department of Population, Policy, and Practice

University College London

14th February 2024

Dedicated to Sebastian Jeans

(- 7th November 2018)

DECLARATION

I, Lucy Karwatowska, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

14th February 2024

ABSTRACT

Disruptive behaviour disorders (DBDs) are a common set of diagnoses affecting around 5.7% of children and adolescents globally. Due to the high personal, social, and economic costs associated with DBDs, there has been a long-standing interest in understanding the environmental, physiological, and genetic factors that may underlie or cause such behaviours. However, until recently, current research on risk factors has been limited due to a lack of generalisability, a failure to account for the complexities of various risk factors, and a reliance on classical correlational analyses, making it hard to establish causation.

This thesis will leverage the recent availability of evidence from large pre-existing datasets, such as genome-wide association studies and registry-based administrative databases. It will use modern causal inference methods to examine putative risk factors for DBDs. The first study in this thesis is a systematic review of research using causal inference methods. The findings highlighted which risk factors have been studied using causal inference methods and which methods have been used to answer causal questions on the development of DBD. The results informed which risk factors and methods were employed in the current thesis.

In the second study, I examined a well-studied environmental risk factor, parenting practices, that could inform preventative interventions for DBDs. I conducted a meta-analysis of 41 studies using causal inference methods. The pooled estimate suggested evidence of a small, causal effect of negative parenting practices on offspring DBDs (r = 0.142; *95% CI* = 0.104, 0.180). I estimated that a hypothetical intervention which reduced negative parenting could lead to a 0.11% reduction in the prevalence of DBDs, the equivalent of 3,614,337 school-aged children worldwide no longer exhibiting clinical levels of DBD symptoms.

In the third study, I investigated whether resting heart rate (RHR), a putative physiological risk factor for disruptive behaviour rarely studied using causal inference methods, is causally related to antisocial behaviour (ASB), a common symptom of

DBDs. Using two-sample Mendelian randomisation (MR) and linkage disequilibrium score regression analyses, I found no evidence of causal associations (B_{IVW} = –0.0004; *95% CI* = -0.004, 0.004, N_{SNPS} = 278) or genetic correlations (r_g = 0.057, *95% CI* = -0.025, 0.139) between RHR and ASB. I discuss how MR can be an effective tool to assess risk factors for DBDs quickly and efficiently, especially for those with a strong genetic basis, provided certain assumptions are met.

In the fourth study, I considered a set of risk factors, exposure to early life adversity, that could help inform public health initiatives for DBDs. I used administrative records from over 1.9 million children and identified four latent trajectory groups of adversity from birth to age six years. I then examined whether group membership was associated with diagnoses of CD or dissocial personality disorder (DPD; *n* = 6,502) or convictions of sexual and violent crimes (*n* = 35,036) before the age of 25 years. The rates of diagnoses and convictions were higher for individuals who experienced early life adversity than individuals in the low adversity group. I also estimated the average treatment effect of a hypothetical intervention that assigned individuals to the lowest adversity group and predicted that it could lead to a two- to three-fold decrease in the probability of diagnoses (males: *ATE* = 2.54, *95% CI* = 2.27, 2.80; females: *ATE* = 3.12, *95% CI* = 2.59, 3.68) and convictions (males: *ATE* = 2.25, *95% CI* = 2.15, 2.35; females: *ATE* = 3.02; *95% CI* = 2.71, 3.39) by the age of 25 years.

This thesis triangulates evidence on risk factors for DBDs by synthesising findings from previous causal inference studies and adopting several novel causal inference methods that rely on different information sources and assumptions.

IMPACT STATEMENT

By leveraging data from large, pre-existing datasets and using novel methods gathered from the modern causal inference literature, this thesis contributes to the field of DBDs in several ways. Firstly, it synthesises existing studies that use causal inference methods to study risk factors for DBDs. Secondly, it examines risk factors that could inform preventative interventions, such as parenting practices, and public health initiatives, such as early life adversity. Thirdly, this thesis demonstrates the utility of more stringent (causal inference) methods that could motivate future research.

Where I uncovered evidence of causal effects, I attempted to estimate the impact on DBDs of hypothetical interventions targeting these risk factors. These findings could interest clinicians and policymakers. My previous work with parents and children with DBDs in the NHS influenced the language used in this thesis. However, I acknowledge that the perspectives of individuals with a lived experience of DBDs are not directly included in the research. Therefore, it is essential to include those perspectives to make a real and lasting impact on the lives of people who need it most.

I presented the findings of my thesis to the wider research community at various conferences, including the Life History Society Conference, where I gave an oral presentation. Moreover, I have published my doctoral work in four peer-reviewed journals as abstracts and scientific articles.

In addition to my doctoral research, I led a successful campaign to change UCL's policy on self-plagiarism. This policy change now allows students to include first-author publications published during their PhD in their thesis without significantly re-writing them. Furthermore, I taught four UCL Doctoral School statistics courses, which helped me improve my knowledge of causal inference while helping other students develop important statistical skills.

LIST OF PUBLICATIONS RELATED TO THIS PHD

Peer-reviewed journal articles

Published

Karwatowska, L., Frach, L., Schoeler, T., Tielbeek, J. J., Murray, J., de Geus, E., Viding, E., & Pingault, J.-B. (2023). Resting heart rate and antisocial behaviour: A Mendelian randomisation study. *Scientific Reports*, *13*(1), 10212. https://doi.org/10.1038/s41598-023-37123-y

Karwatowska, L., Russell, S., Solmi, F., Stavola, B. L. D., Jaffee, S., Pingault, J.-B., & Viding, E. (2020). Risk factors for disruptive behaviours: Protocol for a systematic review and meta-analysis of quasi-experimental evidence. *BMJ Open*, *10*(9), e038258. https://doi.org/10.1136/bmjopen-2020-038258

Under review

Karwatowska, L., Solmi, F., Baldwin, J. R., Jaffee, S., Viding, E., Pingault., J.-B., & De Stavola, B. L. Positive and negative parenting practices, and offspring disruptive behaviour: a meta-analysis of quasi-experimental evidence. (2023). *Under review in Psychological Bulletin.*

In preparation

Karwatowska, L., Viding, E., Pingault., J.-B., Hulvej Rod, N., & De Stavola, B. L. Trajectories of early life adversity and later disruptive behaviour: a nationwide study of over 1.9 million children. (2024). *In preparation for the International Journal of Epidemiology.*

Karwatowska, L., Baldwin, J., Solmi, F., Viding, E., Pingault., J.-B., & De Stavola, B. L. Risk factors for disruptive behaviours: a systematic review and meta-analysis of quasi-experimental evidence. (2024). *In preparation for Psychological Bulletin*.

Conference presentations

Karwatowska, L., Schoeler, T., Viding, E., & Pingault, J.-B. (2019). Evaluating the causal effect of heart rate on antisocial behavior: A multivariable Mendelian randomization study. *Behavior Genetics*, 49(6), 544–544.

Karwatowska, L., Solmi, F., Baldwin, J., Jaffee, S., Viding, E., Pingault, J.-B., & Stavola, B. L. D. (2023). P08 Positive and negative parenting and offspring disruptive behaviour: A meta-analysis of quasi-experimental evidence. *Journal of Epidemiology and Community Health*, *77*, A58–A58. https://doi.org/10.1136/jech-2023-SSMabstracts.117

LIST OF OTHER PUBLICATIONS COMPLETED DURING PHD

Baldwin, J. R., Wang, B., **Karwatowska, L.,** Schoeler, T., Tsaligopoulou, A., Munafò, M. R., & Pingault, J.-B. (2023). Childhood Maltreatment and Mental Health Problems: A Systematic Review and Meta-Analysis of Quasi-Experimental Studies. *American Journal of Psychiatry*, *180*(2), 117–126. https://doi.org/10.1176/appi.ajp.20220174

Lane, C., Hogg, E., **Karwatowska, L. A.**, French, L., Ranieri, V. F., Jesnick, L. G. D., Roberts, C., Scott, S., Senior, R., Skinner, G. C., & Kennedy, E. M. M. (2023). Personalised interventions for subgroups of children with conduct problems. *The Cochrane Database of Systematic Reviews*, *4*(4), CD012746. https://doi.org/10.1002/14651858.CD012746.pub2

Zylbersztejn, A., Lewis, K., Nguyen, V., Matthews, J., Winterburn, I., **Karwatowska, L.**, Barnes, S., Lilliman, M., Saxton, J., Stone, A., Boddy, K., Downs, J., Logan, S., Rahi, J., Black-Hawkins, K., Dearden, L., Ford, T., Harron, K., De Stavola, B., & Gilbert, R. (2023). Evaluation of variation in special educational needs provision and its impact on health and education using administrative records for England: Umbrella protocol for a mixed-methods research programme. *BMJ Open*, *13*(11), e072531. https://doi.org/10.1136/bmjopen-2023-072531

UCL RESEARCH PAPER DECLARATION FORM: referencing the doctoral candidate's own published work(s)

1. For a research manuscript that has already been published (if not yet published, please skip to section 2):

(a) What is the title of the manuscript?

Risk factors for disruptive behaviours: protocol for a systematic review and meta-analysis of quasi-experimental evidence

(b) Please include a link to or doi for the work:

http://dx.doi.org/10.1136/bmjopen-2020-038258

(c) Where was the work published?

The BMJ Open

(d) Who published the work?

British Medical Journal Publishing Group

(e) When was the work published?

2020

(f) List the manuscript's authors in the order they appear on the publication:

Lucy Karwatowska, Simon Russell, Francesca Solmi, Bianca Lucia De Stavola, Sara Jaffee, Jean-Baptiste Pingault, & Essi Viding.

(g) Was the work peer reviewed?

Yes

(h) Have you retained the copyright?

Yes

(i) Was an earlier form of the manuscript uploaded to a preprint server (e.g., medRxiv)? If 'Yes,' please give a link or doi

Yes

If 'No,' please seek permission from the relevant publisher and check the box next to the below statement:

□ I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):

(a) What is the current title of the manuscript?

NA

(b) Has the manuscript been uploaded to a preprint server, e.g., medRxiv?

NA

If 'Yes,' please give a link or doi:

(c) Where is the work intended to be published?

NA

(d) List the manuscript's authors in the intended authorship order:

NA

(e) Stage of publication:

NA

3. For multi-authored work, please give a statement of contribution covering all authors (if single author, please skip to section 4):

LK designed the study and developed the review questions, with guidance from BLDS, J-BP, EV and SRJ. LK drafted and registered the protocol. All authors contributed to revising the protocol manuscript. LK and FS will independently complete the initial screen. LK and SR will independently complete the data extraction and quality assessment. LK will conduct the meta-analyses. LK will write the first manuscript draft and all authors will read and approve the final version of the manuscript. LK will be the guarantor of the review.

4. In which chapter(s) of your thesis can this material be found?

Chapter 2

e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g., if the paper was a single-author work):

Candidate:

Lucy Karwatowska

Date: 29/11/23

Supervisor/Senior Author signature (where appropriate):

Essi Viding

Date: 29/11/23

UCL RESEARCH PAPER DECLARATION FORM: referencing the doctoral candidate's own published work(s)

1. 1. For a research manuscript that has already been published (if not yet published, please skip to section 2):

(a) What is the title of the manuscript?

Resting heart rate and antisocial behaviour: a Mendelian randomisation study

(b) Please include a link to or doi for the work:

https://doi.org/10.1038/s41598-023-37123-y

(c) Where was the work published?

Scientific Reports

(d) Who published the work?

Nature Portfolio

(e) When was the work published?

2023

(f) List the manuscript's authors in the order they appear on the publication:

Lucy Karwatowska, Leonard Frach, Tabea Schoeler, Jorim J. Tielbeek, Joseph Murray, Eco de Geus, Essi Viding & Jean-Baptiste Pingault

(g) Was the work peer reviewed?

Yes

(h) Have you retained the copyright?

Yes

(i) Was an earlier form of the manuscript uploaded to a preprint server (e.g., medRxiv)? If 'Yes,' please give a link or doi

No

If 'No,' please seek permission from the relevant publisher and check the box next to the below statement:

✓ I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):

(a) What is the current title of the manuscript?

NA

(b) Has the manuscript been uploaded to a preprint server e.g., medRxiv?

NA

If 'Yes,' please give a link or doi:

(c) Where is the work intended to be published?

NA

(d) List the manuscript's authors in the intended authorship order:

NA

(e) Stage of publication:

NA

3. For multi-authored work, please give a statement of contribution covering all authors (if single author, please skip to section 4):

L.K. and J.-B.P. contributed to the study conception and design. L.K. conducted the initial analyses and drafted the manuscript. L.F. ran the final analyses. All authors approved the final version of the manuscript for submission.

4. In which chapter(s) of your thesis can this material be found?

Chapter 5

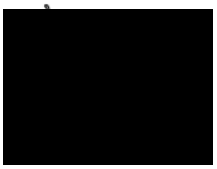
e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g., if the paper was a single-author work):

Candidate:

Lucy Karwatowska

Date: 29/11/23

Supervisor/Senior Author signature (where appropriate):



Jean-Baptiste Pingault

Date: 29/11/2023

UCL RESEARCH PAPER DECLARATION FORM: referencing the doctoral candidate's own published work(s)

1. 1. For a research manuscript that has already been published (if not yet published, please skip to section 2):

(a) What is the title of the manuscript?

NA

(b) Please include a link to or doi for the work:

NA

(c) Where was the work published?

NA

(d) Who published the work?

NA

(e) When was the work published?

NA

(f) List the manuscript's authors in the order they appear on the publication:

NA

(g) Was the work peer reviewed?

NA

(h) Have you retained the copyright?

NA

(i) Was an earlier form of the manuscript uploaded to a preprint server (e.g., medRxiv)? If 'Yes,' please give a link or doi

NA

If 'No,' please seek permission from the relevant publisher and check the box next to the below statement:

□ I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):

(a) What is the current title of the manuscript?

Positive and negative parenting practices and offspring disruptive behaviour: a meta-analysis of quasi-experimental evidence.

(b) Has the manuscript been uploaded to a preprint server e.g., medRxiv?

If 'Yes,' please give a link or doi:

No

(c) Where is the work intended to be published?

Psychological Bulletin

(d) List the manuscript's authors in the intended authorship order:

Lucy Karwatowska, Francesca Solmi, Jessie R. Baldwin, Sara Jaffee, Essi Viding, Jean-Baptiste Pingault & Bianca Lucia De Stavola

(e) Stage of publication:

Under review

3. For multi-authored work, please give a statement of contribution covering all authors (if single author, please skip to section 4):

L.K., F.S., S.J., E.V., J-B.P. & B.L.DS. Participated in the protocol development. L.K. conducted the literature searches. L.K. and F.S. screened the title and abstracts of the articles. L.K., F.S. and B.L.DS screened the full texts of the articles. L.K. and J.R.B. extracted the data from the literature. L.K. conducted the meta-analysis and created all figures and tables. L.K. wrote the first draft and all authors reviewed and commented on subsequent drafts and approved the final version submitted for publication.

4. In which chapter(s) of your thesis can this material be found?

Chapter 4

e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g., if the paper was a single-author work):

Candidate:

Lucy Karwatowska

Date: 29/11/23

Supervisor/Senior Author signature (where appropriate):

Bianca Lucia De Stavola

Date: 01/12/2023

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation and gratitude towards my three supervisors: Prof Bianca Lucia De Stavola, Prof Jean-Baptiste Pingault, and Prof Essi Viding. Each one of my supervisors has played a crucial role in helping me reach this milestone. I would like to extend my special thanks to Prof Essi Viding, whose pioneering work on child mental health inspired me during my undergraduate studies and motivated this research. I am also grateful to Prof Jean-Baptiste Pingault for his confidence in me from our first meeting and for encouraging me to apply to this programme. I am deeply grateful to Prof Bianca Lucia De Stavola, who has been a pillar of support throughout my PhD. Her exceptional intellect, kindness, and guidance have been instrumental in helping me complete this thesis.

I also want to express my heartfelt gratitude to my family, Shân, Jozef, Hannah, and Harriet, who have supported me in countless ways. To my grandparents, Pamela, Doug, Bridget, and Jozef Senior, I am extremely proud to be your granddaughter. To my friends, I cannot possibly express my gratitude to each of you. Clare, Hatty, Janina, Katie, Liz, Lucy, and Serena - thank you for the non-PhD related chats, which encouraged me to take some much-needed space from my thesis. Tif - who, like Essi, has been inspiring me since my undergraduate degree, I am so thankful for your boundless strength, wisdom, and generosity.

To my fellow lab members, Tabea, Kate, and Kai, thank you for your fantastic work and for facilitating many interesting discussions. I am a better researcher for working with you. To the DANLIFE team, particularly Naja, Jessica, Leonie, Megan, and Anna, thank you for welcoming me to Copenhagen and trusting me to analyse your precious data. Doing so changed the course of my research in wonderful ways. To my PhD peers, Gaby, Irina, Jess, Lucie, Ofran and Rebecca, thank you for making this experience infinitely more enjoyable and fulfilling. I am proud to be graduating alongside such intelligent and inspiring women. I am also grateful to another intelligent and inspiring woman, Dr Jessie Baldwin, my friend and academic mentor,

for her positivity and guidance during my PhD. Her passion for research helped to reignite my own when I needed it.

Finally, I would like to express my huge gratitude to my examiners, Prof Barbara Maughan and Prof Glyn Lewis, for their flexibility, understanding and interest in not only my PhD but also my experiences.

I dedicate this thesis to Sebastian, whose memory continues to profoundly impact me. I will think of you often and remember you always.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Terms

ADHD	Attention Deficit Hyperactivity Disorder
ANS	Autonomic Nervous System
ASB	Antisocial Behaviour
ASD	Autism Spectrum Disorders
ASPD	Antisocial Personality Disorder
ATE	Average Treatment Effect
BIC	Bayesian Information Criterion
CAMHS	Child and Adolescent Mental Health Services
CD	Conduct Disorder
CI	Confidence Interval
CP	Conduct Problems
CPR	Civil Personal Registration
DAG	Directed Acyclic Graph
DBD	Disruptive Behaviour Disorder
DPD	Dissocial Personality Disorder
DSM	Diagnostic and Statistical Manual of Mental Disorders
DZ	Dizygotic
ECG	Electrocardiogram
ES	Effect size
EXT	Externalising Symptoms
GBTM	Group-Based Multi-Trajectory Modelling
GWAS	Genome Wide Association Study
HIC	High Income Country
HPA	Hypothalamic-Pituitary-Adrenal
HR	Hazard Ratio
HRV	Heart Rate Variability
ICD	International Classification of Diseases
INFO	Information Score
INT	Internalising Symptoms
IPWT	Inverse Probability of Treatment Weighting
IV	Instrumental Variable
IVF	In Vitro Fertilisation
IVW	Inverse-Variance Weighted
LA	Low Adversity

LD	Linkage Disequilibrium
LDSC	Linkage Disequilibrium Score Regression
LFD	Loss and Family Dynamics
LMIC	Low and Middle Income Country
MD	Material Deprivation
MOOSE	Meta-analysis of Observational Studies in Epidemiology
MR	Mendelian Randomisation
MRI	Magnetic Resonance Imaging
MZ	Monozygotic
OCC	Odds of Correct Classification
ODD	Oppositional Defiant Disorder
PMT	Parent Management Training
PRESSO	Pleiotropy RESidual Sum and Outlier
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International Prospective Register of Systematic Reviews
pvRSA/HF	Peak-Valley Respiratory Sinus Arrhythmia or High-frequency Power
RAPS	Robust Adjusted Profile Score
RCT	Randomised Control Trials
rGE	Gene-Environment Correlation
RHR	Resting Heart Rate
RMSSD	Root Mean Square of the Successive Differences of Inter-beat Intervals
RR	Risk Ratio
SD	Standard Deviation
SDNN	Standard Deviation of the Normal-to-Normal Inter-beat Intervals
SDQ	Strength and Difficulties Questionnaire
SE	Standard Error
SNP	Single Nucleotide Polymorphism
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TTE	Target Trial Emulation
UK	United Kingdom
USA	United States of America

Cohorts

AddHealth	The National Longitudinal Study of Adolescent to Adult Health
ALSPAC	The Avon Longitudinal Study of Parents and Children
ARIC	Atherosclerosis Risk in Communities Study

ATR	Australian Twin Register
BeTwiSt	Beijing Twin Study
BIG	Brain Imaging Genetics
BroadABC	Broad Antisocial Behavior Consortium
BYS	Boricua Youth Study
C-IVF	Cardiff IVF study
CHDS	Christchurch Health and Development Study
CHS	Cardiovascular Health Study
CNLSY	Children of the National Longitudinal Survey of Youth
COGA	Collaborative Study on the Genetics of Alcoholism
CPP	Collaborative Perinatal Project
DANLIFE	DANish LIFE Course Cohort
DMTS	Danish Mother of Twins Survey
DNBC	Danish National Birth Cohort
E-Risk	Environmental Risk Longitudinal Twin Study
ECLS-B	Early Childhood Longitudinal Study, Birth Cohort
ECLS-K	Early Childhood Longitudinal Study - Kindergarten Cohort
EDEN	Etude des Déterminants du développement et de la santé de l'ENfant
EGDS	Early Growth and Development Study
ELPI	Encuesta Longitudinal de la Primera Infancia Cohort
ERA	English and Romanian Adoptees Study
FFCWS	Fragile Families and Child Wellbeing Study
FHS	Framingham Heart Study
FINGESTURE	FINnish GEnetic STUdy of aRrhythmic Events
FinnTwin	Finnish Twin Cohort
FLP	Family Life Project
GenR	Generation R Study
GSA	The Genetics of Sexuality and Aggression
GSMS	Great Smoky Mountains Study
GTR	Groningen Twin Registry
GUI	Growing Up in Ireland Child Cohort
HBHC	Healthy Babies Healthy Children Study
HSIS	Head Start Impact Study
iPSYCH	Lundbeck Foundation Initiative for Integrative Psychiatric Research
J-SHINE	Japanese Study of Stratification, Health, Income, and Neighborhood
KORA S4	KOoperative gesundheitsforschung in der Region
LEMENGHO-	FLEMish study on Environment, Genes and Health Outcomes
EPOGH	– European Project on Genes in Hypertension

LIST	Longitudinal Israeli Study of Twins
LSAC	Longitudinal Study of Australian Children
MCS	Millennium Cohort Study
MCTFR	Minnesota Center for Twin and Family Research
MESA	Multi-Ethnic Study of Atherosclerosis
MIDS	Midwest Infant Development Study
MO-MATCH	Missouri Mothers and Their Children Study
MOAFTS	Missouri Adolescent Female Twin Study
MoBa	Norwegian Mother, Father and Child Cohort Study
MTFS	Minnesota Twin Family Study
N2CAP	Northeast-Northwest Collaborative Adoption Projects
NCS-A	National Comorbidity Survey Replication Adolescent Supplement
NEAD	Nonshared Environment and Adolescent Development Project
NESDA	Netherlands Study of Depression and Anxiety
NICHD SECCYD	National Institute of Child Health and Human Development Study of Early Child Care and Youth Development
NLCSY	National Longitudinal Survey of Children and Youth
NLSY79	National Longitudinal Study of Youth
NSCAW II	National Survey of Child and Adolescent Well-Being
NSCH	National Survey of Children's Health
NTR	Netherlands Twin Register
OCHS	Ontario Child Health Study
OZALC	Australian Twin-Family Study of Alcohol Use Disorder
PAGES	Phenomics and Genomics Sample
PGS	Pittsburgh Girls Study
PHDCN	Project on Human Development in Chicago Neighborhoods
PIVUS	Prospective Investigation of the Vasculature in Uppsala Seniors
PREVEND	Prevention of Renal and Vascular ENd- stage Disease
PSID CDS	Panel Study of Income Dynamics Study Child Development Supplement
PYS	Pittsburgh Youth Study
QIMR	QIMR Berghofer Medical Research Institute
QLSCD	Quebec Longitudinal Study of Child Development
RSI+2	Rotterdam Study
S4S	Spit for Science
SBS	Sisters and Brothers Study
SIBS	Sibling Interaction and Behavior Study
TAMBAHS	Twins and Multiple Births Association Heritability Study
TBED-C	Twin Study of Behavioral and Emotional Development in Children

TCHAD	Twin Study of Child and Adolescent Development Study
TEDS	Twins Early Development Study
TESS	Trondheim Early Secure Study
TFaB	The Twins, Family and Behaviour Study
TOSS	Twin and Offspring Study in Sweden
TRAILS	Tracking Adolescents' Individual Lives Survey
ULSAM	Uppsala Longitudinal Study of Adult Men
VATSPSUD	Virginia Adult Twin Study of Psychiatric and Substance Use Disorders
VET	Vietnam Era Twin Registry
VgHRV	Genetic Variance in Heart Rate Variability Consortium
WACS	Wales Adoption Cohort Study
WES	Women's Employment Study
YFS	Cardiovascular Risk in Young Finns Study
YPCA	Young Parents and Their Children in Australia Study
Z-PROSO	Zurich Project on the Social Development of Children and Youths

1 INTRODUCTION

Chapter overview

In this Chapter, I will define the key concepts covered in this thesis. I will start by discussing how disruptive behaviour disorders are defined in the literature and how I chose to define them in this thesis. I will then review the current evidence on risk factors for disruptive behaviour. Then, I will describe some common threats to internal and external validity in observational studies. Finally, I will present the main concepts invoked when aiming to infer causality from observational studies and outline some common methods that can strengthen causal inference in epidemiology.

1.1 Background

Individuals who display disruptive behaviour disorders (DBDs) in childhood (also referred to as externalising behaviour [EXT], conduct problems [CP], conduct disorders [CD] and oppositional defiant disorder [ODD]) engage in a range of repetitive and troublesome behaviours. The International Statistical Classification of Diseases and Related Health Problems: Tenth Revision (ICD-10; World Health Organization, 2004) defines CD (F91) as *"characterised by a repetitive and persistent pattern of dissocial, aggressive, or defiant conduct"*. Examples of such behaviours include *"excessive levels of fighting or bullying, cruelty to other people or animals, severe destructiveness to property, fire-setting, stealing, repeated lying, truancy from school and*

running away from home, unusually frequent and severe temper tantrums, and disobedience."

DBDs can be categorised into different subgroups based on several factors such as age-of-onset (e.g., childhood- versus adolescent-onset; Moffitt, 1993), type and severity of disruptive behaviour symptoms (e.g., subclinical CP versus clinical CD) and the presence/absence of limited prosocial emotions (i.e., deficits in empathy and remorse; Frick et al., 2003). Additionally, DBDs can co-occur with other conditions such as internalising [INT] disorders (e.g., anxiety and depression; Polier et al., 2012), other EXT disorders (e.g., attention deficit hyperactivity disorder [ADHD; Frick & Ellis, 1999) and developmental disorders (e.g., autism spectrum disorders [ASD; Carter Leno et al., 2021). The heterogeneity of DBDs is well-documented (Viding & McCrory, 2020) and will not be the focus of this thesis. Instead, I will define DBDs as clinical diagnoses (e.g., CD, ODD and antisocial personality disorder [ASPD]) or symptoms (e.g., antisocial or violent behaviour) associated with broadly defined disruptive behaviour.

In 2015, estimates suggested that around 5.7% or approximately 113 million children and adolescents globally exhibit symptoms of DBDs, either ODD or CD (Polanczyk et al., 2015). This figure is similar in the United Kingdom (UK), with an estimated 6% of individuals between 5 and 15 years old displaying either CD or ODD (Green et al., 2005; Pilling et al., 2013). DBDs are the most common reason for referral to Child and Adolescent Mental Health Services (CAMHS), surpassing other common childhood mental health diagnoses such as ADHD and INT disorders (Green et al., 2005).

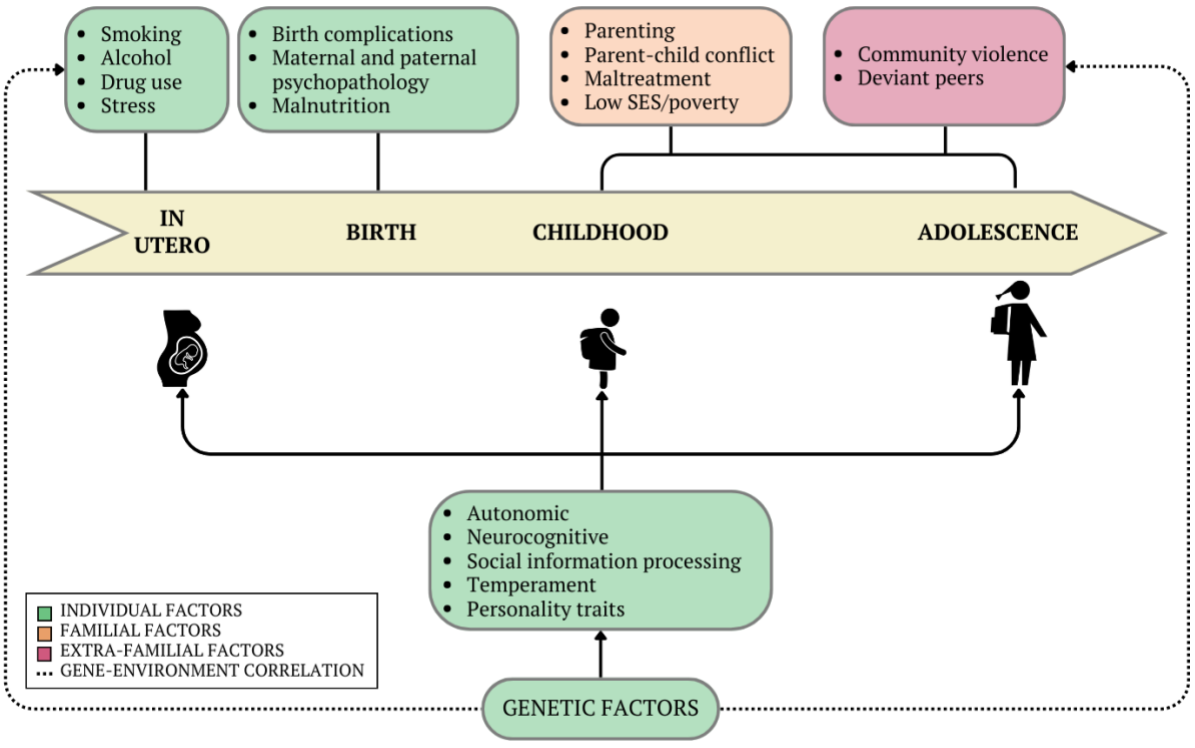
Although DBDs typically start in childhood, their impact extends well beyond this developmental stage. For some individuals, DBDs demonstrate continuity over the

life course, with up to 50% of individuals who display CD in childhood continuing to exhibit such behaviour into adolescence and adulthood (E. D. Barker & Maughan, 2009; Maughan & Kim-Cohen, 2005; Moffitt, 1993; Simonoff et al., 2004). This continuity often results in a diagnosis of ASPD (Diagnostic and Statistical Manual [DSM]; (American Psychiatric Association, 2013) or dissocial personality disorder (ICD; W. E. Copeland et al., 2009; Moffitt et al., 2008). DBDs are associated with many negative outcomes throughout life, making them highly impairing disorders. Individuals diagnosed with DBDs during childhood are at higher risk of experiencing other poor mental health outcomes, such as substance misuse, early mortality due to increased risk of injury, aggressive and criminal behaviour leading to convictions, poorer educational outcomes, and unemployment (Bevilacqua et al., 2018; Burt et al., 2018; Fairchild et al., 2019; Huesmann et al., 2009; Piquero et al., 2007; Colman et al., 2009).

Indeed, individuals with DBDs engage with far more criminal justice, health, and social welfare services than the population average (Rivenbark et al., 2018). As such, DBDs pose a considerable personal, social, and economic burden on individuals and society (Erskine et al., 2014; Heeks et al., 2018; Rivenbark et al., 2018). The cost of crime in the UK is estimated at £50 billion annually. Sexual and violent crimes, e.g. homicide, rape, and violence with injury, contribute the largest proportion to these costs because of the physical and emotional costs associated with such crimes (Heeks et al., 2018). Furthermore, DBDs are responsible for 12 times more "years lived with disability," a measure of disease burden, than ADHD (Erskine et al., 2014). Despite this, ADHD receives more research funding than DBDs (MQ Transforming Mental Health., 2021). However, DBDs are a significant public health concern, and research identifying causal risk factors as early as possible is crucial to inform interventions that aim to prevent DBDs and reduce the associated poor long-term outcomes.

1.2 Current understanding of *putative* risk factors

The development of DBDs is understood to be influenced by a combination of genetic and environmental risk factors and their interactions (see Figure **1.1**). Behavioural-genetic studies (e.g., twin and adoption studies) suggest that DBDs are 50% heritable (Cadoret & Stewart, 1991; Lewis & Plomin, 2015; Mason & Frick, 1994; Miles & Carey, 1997; Moffitt, 2005; Polderman et al., 2015; Rhee & Waldman, 2002; Salvatore & Dick, 2018), which indicates that a substantial proportion of the population variation in DBDs is due to environmental influences (Beaver et al., 2018). Environmental factors can represent crucial targets for early intervention and prevention efforts as, in theory, many can be modified, and if they are truly causal, this will lead to a reduction in DBD symptoms (Derzon, 2010; Hawkins et al., 2000; Jaffee, Strait, et al., 2012; Murray & Farrington, 2010).



Adapted from Fairchild, G., Hawes, D. J., Frick, P. J., Copeland, W. E., Odgers, C. L., Franke, B., Freitag, C. M., & Brito, S. A. D. (2019). Conduct disorder. Nature Reviews Disease Primers, 5(1), 43. https://doi.org/10.1038/s41572-019-0095-y

Figure 1.1 *Overview of putative risk factors associated with disruptive behaviour disorders throughout the life course.*

Figure **1.1** provides an overview of the risk factors suggested to be causally related to DBD in the current literature. These risk factors can be categorised into individual, familial and extra-familial factors.

1.2.1 Individual level factors

Many individual-level factors have been posited to play a role in the aetiology of DBD. These can be further categorised into genetic, pre- and perinatal, hormonal, neural, psychophysiological, and cognitive risk factors.

1.2.1.1 Genetic factors

Although the heritability estimates for DBDs of 50% are not as high as for other developmental disorders, such as ADHD (74% heritability; Faraone & Larsson, 2019) or ASD (80% heritability; Tick et al., 2016), it still indicates that genetic vulnerabilities play a significant role in the development of DBDs. If genetic factors confound the relationship between the exposure of interest and outcome (i.e. are associated with both the exposure and outcome variables and not downstream from the exposure; VanderWeele, 2019) then researchers need to control for genetic factors in the study design or analyses (see Section 1.3.1.1 below).

1.2.1.2 Pre- and perinatal factors

The prenatal environment is vital in offspring development (D. J. P. Barker, 1990). Various prenatal factors have been suggested to be possible risk factors for DBDs, including in-utero exposure to toxins (e.g. tobacco, alcohol, and caffeine; Haan et al., 2022) and stress hormones (e.g. cortisol related to stress and depressive symptoms; MacKinnon et al., 2018). Perinatal factors, including obstetric complications and low birth weight, have also been associated with offspring DBDs (Hazebroek et al., 2019). It is thought that these prenatal and perinatal factors are distal factors which increase

DBDs through their influence on other factors, such as low IQ and low academic achievement (L. Ellis & Walsh, 2003; Liu et al., 2009; Moffitt, 2005; Pennington & Ozonoff, 1996; Ttofi et al., 2016).

1.2.1.3 Hormonal factors

The hormone most linked to DBDs, especially violence and aggression, is testosterone (Book et al., 2001), specifically the imbalance between testosterone and other hormones such as serotonin and cortisol. It is thought that higher testosterone levels in males may explain the sex differences in the prevalence of DBDs (Pavlov et al., 2012).

1.2.1.4 Neural factors

Resting-state functional Magnetic Resonance Imaging (fMRI) and structural MRI studies have found atypical neural responses, brain structure and structural connectivity in individuals with DBDs in areas of the brain related to emotion processing (e.g. the amygdala), decision-making (e.g. the prefrontal cortex) and threat response (e.g. the amygdala, prefrontal cortex and the anterior cingulate cortex; De Brito et al., 2021; Fairchild et al., 2019; Viding et al., 2023).

1.2.1.5 Cognitive-affective factors

The differences in neural function and structure of individuals with DBDs are thought to result in neurocognitive impairments such as difficulty in recognising emotions, understanding empathy, making decisions, and processing reward and punishment cues (see Fairchild et al., 2019).

1.2.1.6 Psychophysiological factors

Markers of general emotional arousal, such as skin conductance and heart rate, are commonly associated with DBDs. As discussed further in Chapter 5, researchers suggest that low resting heart rate is the *"best-replicated biological correlate"* (Ortiz & Raine, 2004, p. 154) and a *"possible causal risk factor"* (Portnoy & Farrington, 2015, p. 42) for antisocial behaviour, a common symptom of DBDs.

1.2.2 Familial factors

Factors that are shared between family members also have an impact on the development of DBDs.

1.2.2.1 Adverse childhood experiences (ACEs)

Most familial factors posited to play a role in DBDs are included in the umbrella term adverse childhood experiences (ACEs; Felitti et al., 1998). ACEs are experiences that occur outside of a child's expected environment and are significant enough to require psychological, social, and neurodevelopmental adaptation by the individual (E. McCrory et al., 2012; E. McCrory & Viding, 2015; McLaughlin et al., 2012; The Lancet Public Health, 2021). ACEs are consistently linked to the development of various childhood mental health conditions, including DBDs. Examples of ACEs include maltreatment, neglect, physical and emotional abuse (E. McCrory et al., 2012; E. McCrory & Viding, 2015; McLaughlin et al., 2012; McLaughlin & Lambert, 2017; Wilson et al., 2009); witnessing high levels of parental conflict or intimate partner violence (Harden, Turkheimer, et al., 2007; Meyer et al., 2000); parental separation (Fergusson et al., 1994); death of a parent (Berg et al., 2019), having a parent with high levels of psychopathology, such as internalising (Hay et al., 2010) or externalising symptoms (Frick et al., 1992).

1.2.2.2 Parenting practices

Caregivers play a crucial role in children's development through social learning (Bandura & Walters, 1963) and operant conditioning (Skinner, 1950). While positive parenting practices, such as warm, supportive parenting, have been associated with a reduction in offspring DBDs, negative parenting practices, such as negative expressed emotion and harsh or inconsistent parenting, have been suggested to increase DBDs (Cooke et al., 2022; Mingebach et al., 2018; Pinquart, 2017; Rothbaum & Weisz, 1994).

1.2.2.3 Low family income

Globally, 1 in 6 children experience extreme poverty (Silwal et al., 2020), with 30% of children in the UK estimated to live in poverty (Marmot, 2020; Marmot et al., 2008). Individuals who experience poverty in early life have a greater risk of developing DBD symptoms (Jaffee, Hanscombe, et al., 2012; Walker et al., 2011). Furthermore, research has shown that cash transfers may increase mental health and well-being (McGuire et al., 2022; Thomson et al., 2022; Zimmerman et al., 2021) and specifically decrease levels of disruptive behaviour and crime (Akee et al., 2010; Costello et al., 2003).

1.2.3 Extra-familial factors

Although not directly examined in this thesis, distal factors that function at the community and societal level have also been suggested to influence disruptive behaviour.

1.2.3.1 Peer problems, bullying victimisation, and high delinquency peers

Outside of the family, associating with highly delinquent peers and being a victim of bullying has been linked to an increased risk of DBDs (Schoeler et al., 2018; Singham et al., 2017; Thornberry et al., 1994; Warr, 1993). Like many of the risk factors

discussed already, the relationship between exposure to bullying victimisation, socialising with high-delinquency peers, and exhibiting DBD symptoms is complex and could be bi-directional (Jaffee, Strait, et al., 2012; Thornberry et al., 1994).

1.2.3.2 High-crime, high-poverty neighbourhoods

Individuals who live in high-crime, high-poverty neighbourhoods also show an increased risk of DBDs (Leventhal & Brooks-Gunn, 2000; Odgers et al., 2012).

1.2.4 Overview of potential risk factors

Importantly, these individual, familial, and extra-familial risk factors are interrelated, making it difficult to determine cause and effect. Indeed, individuals who experience one risk factor are much more likely to experience others (Ben-Shlomo & Kuh, 2002; Dong et al., 2004). Examining risk factors in isolation without considering the effect of other related factors may bias the association between the putative risk factor of interest and DBDs. For example, research using methods that are more robust to genetic confounding has suggested that the relationship between many of these risk factors, such as pre- and perinatal factors (Haan et al., 2022), and DBDs is confounded by shared genetics. Furthermore, longitudinal research using informant-reported measures of parenting practices and/or DBDs has suggested that previous associations between certain risk factors, such as parenting practices, and offspring DBDs may be biased by reverse causation and shared method variance (see Section 1.3 below for definitions; Jaffee et al., 2012; Murray et al., 2018; Thornberry et al., 1994).

1.3 Threats to the validity of extant findings

The randomisation of individuals to intervention groups enables RCTs to examine causal effects in situations where the trial is double-blind and does not suffer from informative loss to follow-up. The presence of confounders (i.e. variables that are associated with both the exposure and outcome variables and are not downstream from the exposure, discussed in detail below) can bias results. Randomisation guarantees that potential confounders are similarly distributed in treatment and control groups, and for this reason, RCTs are often touted as the "gold standard" for causal inference. However, RCTs are expensive to run and frequently have a short follow-up duration, difficulty in recruiting a broad range of individuals, and different levels of attrition by intervention arm (Bärnighausen, Tugwell, et al., 2017; Hernán & Robins, 2016). Therefore, although RCTs have high internal validity (i.e., can assess causal effects), they can have very low external validity (i.e., cannot be generalised to other situations or populations), as they are often conducted using specific populations (i.e. at-risk children; (Jaffee, Strait, et al., 2012). Furthermore, RCTs are often infeasible and unethical to assess many of the abovementioned putative factors. For example, you cannot randomise a child to experience childhood adversity. Finally, RCTs evaluate the effect of interventions that influence specific risk factors, but it is often unclear what part of the intervention is effective (Leijten et al., 2022).

When available, researchers can use observational data to investigate causal effects. However, common threats to the validity of findings in observational research include confounding, selection bias and measurement bias. These interrelated concepts pose challenges when assessing causality from existing research, which will be the subject of the following section.

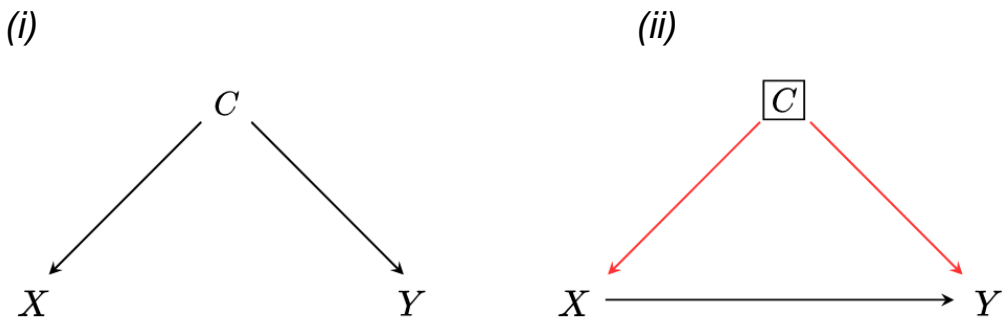
1.3.1 Threats to internal validity

Internal validity refers to the degree to which a research study can estimate a causal relationship between the potential risk factor and outcome in the population targeted by the study. In the following sections, I will discuss how common examples of confounding biases in epidemiological research threaten the internal validity of extant findings.

1.3.1.1 Confounding

A confounder is a variable associated with or a common cause of the exposure and outcome variables that is not downstream from the exposure (VanderWeele, 2019). If a confounder is not controlled for in either the study design or analyses, it can generate a non-causal association between the factor of interest and the outcome, thus distorting an existing causal association or creating the appearance of one when none existed, resulting in bias.

Directed acyclic graphs are path diagrams (Pearl, 1995) that enable researchers to visualise the key assumptions made *a-priori* when investigating putative causal relationships, including the presence of confounding variables (Greenland et al., 1999). A simple example of a setting where the relationship between an exposure (*X*) and an outcome (*Y*) is affected by a confounder (*C*) is shown in Figure 1.2. The pathway between the exposure and outcome via the confounder is often described as a "backdoor path" (Pearl, 2009). If this backdoor path is not blocked, then it can appear that *X causes Y* when, in fact, they are merely associated *through C*.



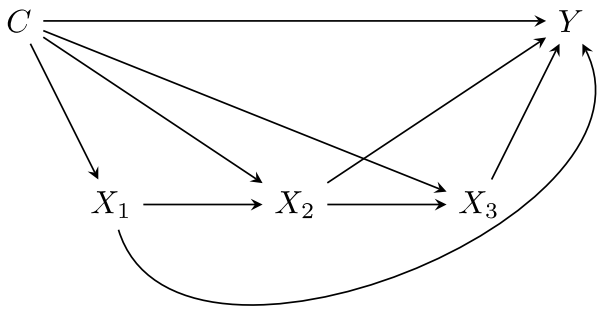
Note. X = exposure, Y = outcome, C = confounder. ***i)*** *a "backdoor path" from X to Y through C.* ***ii)*** *controlling for the confounder allows examination of causal pathways between X and Y.*

Figure 1.2 *A simplified directed acyclic graph of a setting affected by confounding.*

When considering confounding, it is useful to discuss two distinct settings when the exposure of interest is *(a)* time-fixed ("single-point") or *(b)* time-varying. In the former setting, confounding may arise from baseline confounding, and techniques such as regression adjustment, matching or stratification by the identified confounders may be sufficient to block any backdoor paths between the exposure and the outcome. In the latter setting, confounding may arise from baseline and time-varying confounding, and therefore, more appropriate methods, known as g-methods, are required (see Section 1.4.2.2 below).

Baseline confounding, also known as time-fixed confounding, occurs when baseline characteristics (e.g., sex, birth weight) are associated with the outcome of interest and an individual's exposure to a potential risk factor. Baseline confounding can occur in settings with time-fixed *and* time-varying exposures. As I will discuss further in Chapter 4, when investigating the potential causal effect of parenting practices on DBDs, a potential baseline confounder of this relationship may be child sex, whereby whether a child is born male or female may illicit different parenting practices from their parents and may influence the child's propensity to engage in disruptive behaviour. A DAG for a setting with a baseline confounder, which influences the

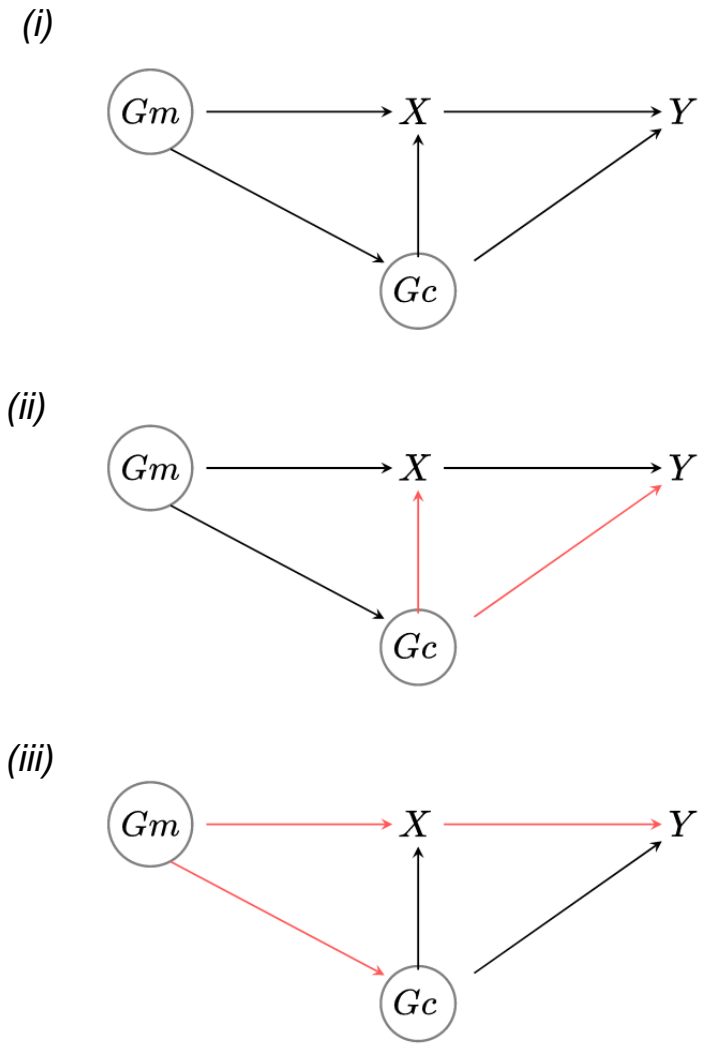
relationship between a time-varying exposure and time-varying outcomes, is shown in Figure 1.3.



Note. X = exposure, Y = outcome, C = confounder, X_n denotes the value of X at time n.

Figure 1.3 *A simplified directed acyclic graph of a setting with a time-varying exposure that is affected by baseline confounding.*

A special case of time-fixed confounding is genetic confounding, which occurs when genetic factors are associated with both the exposure and the outcome (Pingault et al., 2018). For instance, genes influencing parenting practices could be shared with the child and affect the child's disruptive behaviour. In Figure 1.4 *(i),* the genetic variants in the mother (G_m) that influence the environmental exposure (parenting practices; *X*) are shared with the child (G_c) and these genetic variants in the child are associated with the outcome, disruptive behaviour (*Y*).



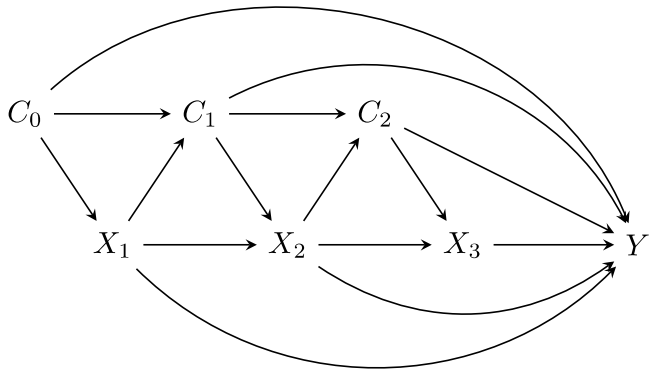
Note. X = exposure, Y = outcome, G_m genes of the mother, G_c = genes of the child.

Figure 1.4 *A simplified directed acyclic graph of settings affected by (i) gene-environment correlations, with (ii) active gene-environment and (iii) passive gene-environment correction.*

Genetic confounding occurs through passive and active (or evocative) gene-environment correlation (r_{GE}). Passive r_{GE} occurs when children inherit genes from their parents that influence their parentally driven environment (Figure 1.4 *(ii)*; backdoor path from $G_c \leftarrow G_m \rightarrow X$). For example, a parent's genotype may affect their parenting practices. If the child inherits these same genes, the child's genotype will be correlated with their parentally driven environment, in this case, the parenting practices. Active r_{GE} occurs when children inherit genes from their parents that influence their self-driven environment (Figure 1.4 *(iii)*; $G_c \rightarrow X$). For instance, an

individual inherits the genes for disruptive behaviour, and this behaviour elicits more negative parenting practices from their parents. Genetically informed causal inference methods, which I describe in more detail in Section 1.4.2.1.3 and are included in my meta-analysis in Chapter 4, can be used to control for r_{GE}.

Time-varying confounding occurs in longitudinal settings when a third variable, which is downstream from the exposure, influences both the outcome of interest and future levels of the exposure. In Figure 1.5, parental mental health at baseline (C_0) may influence later parental mental health (C_1) and also affect later family income (X_1) contributing to disruptive behaviour (Y). This creates the situation where, by being affected by the exposure at an earlier time point (X_1) and by influencing future exposure values (X_2), the third variable (C_1) is both a confounder of the exposure-outcome relationship ($X_2 \leftarrow C_1 \rightarrow Y$) and an intermediate variable (i.e., on the causal pathway) between the earlier exposure and outcome ($X_1 \rightarrow C_1 \rightarrow Y$). G-methods, which I describe in Section 1.4.2.2 and use in Chapter 6, can account for the effects of time-varying confounding.

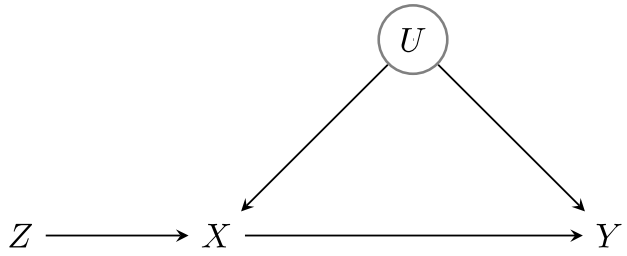


Note. X = exposure, Y = outcome, C = confounder, X_n denotes the value of X at time n.

Figure 1.5 *A simplified directed acyclic graph of a setting with a time-varying exposure that is affected by time-varying confounding.*

1.3.1.1.1 Lack of randomisation of the exposure

Randomisation of the exposure ensures that any missing data or (unmeasured) confounding is independent of the exposure and outcome. Therefore, individuals who receive the exposure and those who do not are similar to or exchangeable with one another. This enables experimental studies (i.e., RCTs) to estimate the causal effect of assignment to the intervention (see Section 1.4 below for more detail). Figure 1.6 depicts this setting where *Z* is randomisation, which influences the uptake of an intervention, *X*. Randomisation (*Z)* is said to be an instrumental variable for *X*. Other instrumental variable approaches, such as natural experiments and Mendelian randomisation (MR), take advantage of naturally occurring events or situations, such as policy initiatives or genetic variants, which also separate the pathways from exposures and other variables. I describe these different approaches in Section 1.4.2.1.2 and use MR in Chapter 5.



Note. Z = instrumental variable, X = exposure, Y = outcome, U = unmeasured confounders.

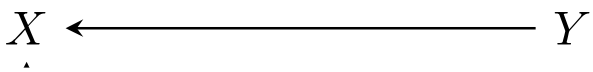
Figure 1.6 *A simplified directed acyclic graph of a setting in which there is an instrumental variable.*

1.3.1.1.2 Reverse causation

Although there may be a causal relationship between factors, the direction of effects may be unclear. In other words, it might be difficult to determine cause and effect. For example, as well as influencing disruptive behaviour, symptoms of parental psychopathology could also be a consequence of children's DBDs. Reverse causation can bias results if the estimate of the effect of *X* on *Y* reflects the impact of *Y* on X (see Figure 1.7). Reverse causation is particularly likely when variables are time-

varying and influence each other. Understanding the temporal ordering of variables is, therefore, crucial to identifying true causal effects. If the complexities between the variables are not acknowledged (e.g., through DAGs) and/or adequately accounted for, the associations found may be biased.

1.3.1.1.3 Collider bias



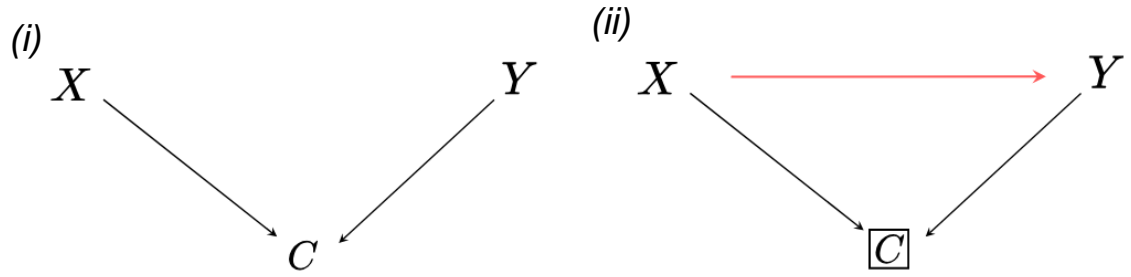
Note. X = exposure, Y = outcome.

Figure 1.7 *A simplified directed acyclic graph of a setting affected by reverse causation.*

Collider bias is another source of bias in observational research (Figure 1.8). Collider bias occurs when the study design or analysis conditions on a third variable (by restriction, stratification, or regression adjustment) that is affected by the exposure of interest *and* by the outcome (i.e. a variable that is downstream from the exposure and shares causes with the outcome; Cole et al., 2010; Figure 1.8, *(i)*). Conditioning on a collider induces an additional (non-causal) association between exposure and outcome via a new link among its "parents" (Figure 1.8, *(ii)*) and thus biases estimates of causal effects of the exposure on the outcome.

One common source of collider bias in epidemiology is selection bias/lack of representativeness in the sample (discussed below in Section 1.3.2.1 on selection bias). Collider bias is induced when a sample is restricted either due to selection bias into the study, attrition over time or restricting the analyses to complete records (Munafò et al., 2018). If the missing data mechanism is driven by the exposure and other factors related to the outcome, any associations between the exposure and the outcome will be impacted. One example of this is that individuals who experience adversity (exposure) and also individuals who exhibit disruptive behaviour (outcome)

are less likely to participate and more likely to drop out of research studies (Doidge et al., 2017); this selection bias could induce spurious associations between exposure and outcome. One way to avoid conditioning on colliders (e.g. sample characteristics) is by using population-based data, such as national registers, with limited missing data (Chapter 6).



Note. X = exposure, Y = outcome, C = collider.

Figure 1.8 *A simplified directed acyclic graph of a setting in which there is a collider on the path between X and Y.*

1.3.1.2 Misclassification and measurement error bias

Misclassification and measurement error bias refers to the impact of a measurement tool or test that inaccurately classifies or measures individuals in terms of their exposures, outcomes, or confounders. For example, if exposed/non-exposed participants are incorrectly classified as non-exposed/exposed and vice versa (K. T. Copeland et al., 1977). Although misclassification can be produced by random error, and, therefore, a degree of measurement error is unavoidable, much of the research on psychopathology relies on subjective questionnaires and retrospective recall, which increases the chance of misclassification. Two common biases affecting research based on questionnaires are shared method and recall bias.

1.3.1.2.1 Shared method bias

Shared method bias, also referred to as "common method bias," is defined as "the variance that is attributable to the measurement method rather than to the constructs the measures represent" (Podsakoff et al., 2003, p. 879). Shared method variance is an artefactual covariance between the exposure and outcome if the same informant reports both measures (De Los Reyes et al., 2009). Research on other mental health diagnoses suggests that shared method variance may account for the magnitude of reported effects in studies that use the same informant for the exposure and the outcome (Francis et al., 2023; Schoeler et al., 2018). I will specifically examine shared method bias in my meta-analysis in Chapter 4.

1.3.1.2.2 Recall bias

Recall bias can occur when exposure and/or outcomes are retrospectively measured, and their reporting may be influenced by current exposure and/or outcomes. Some examples of recall bias include events occurring in early childhood being more likely to be forgotten due to infantile amnesia (Travaglia et al., 2016) or "recency" effects where the accuracy of event recall decreases as the time elapsed between the event and recalling the event increases (Hänninen & Soininen, 1997). It is also possible that individuals who have higher levels of disruptive behaviour may be more likely to retrospectively report certain risk factors (e.g., exposure to negative parenting practices) due to negative cognitive biases (Beck, 2008) and/or exposure suspicion bias (Sackett, 1979). Likewise, parents of offspring with DBD symptoms might incorrectly recall risk factors (e.g., difficult temperament or association with delinquent peers) to explain their offspring's behaviour (Bower, 1981; Brewin et al., 1993; A. Reuben et al., 2016) or they may underreport exposures/outcomes due to feelings of embarrassment (Sackett, 1979).

Prospectively recorded objective measures can aid the classification of individuals. Researchers can achieve this is by using genetic instruments (Chapter 5), longitudinal data (Chapters 4 and 6), and administrative data available from registries (Chapter 6), as these are less susceptible to measurement error and reporting bias.

1.3.2 Threats to external validity

The term 'external validity' refers to the extent to which the findings of a study can be generalised to other populations, settings, and time periods beyond the specific context in which the study was conducted.

1.3.2.1 Selection bias

The main reason for failing external validity is selection bias. As I discussed in Section 1.3.1.1.3 on collider bias, selection bias can occur at any stage of a research study (i.e. recruitment, inclusion, follow-up, and analyses). Selection bias means that the study population may not represent the target population and that the causal relationship of interest is not transportable to the target population because of heterogeneity of effect across different populations (Richiardi et al., 2013; Rothman et al., 2013). In terms of representativeness, or lack thereof, many studies that were established to investigate the development and stability of DBDs over the life course have historically included small and/or unrepresentative samples. Results obtained on selected sub-populations, such as at-risk or clinical samples, are not generalisable to the wider population with less severe symptoms of DBDs. Furthermore, research on small samples often cannot sufficiently investigate the complex relationships (i.e., heterogeneity of effects) between putative risk factors, confounders, and outcomes due to lack of power and incomplete information. Therefore, the utility of evidence generated by these studies is limited.

In summary, existing research outputs are limited in several ways. Many studies suffer from unaccounted confounding and the impact of measurement/misclassification error bias. Furthermore, until relatively recently, existing research has been limited by a lack of generalisability, a failure to acknowledge and account for the complexities of the many exposures, a reliance on classical correlational analyses and, therefore, an inability to determine causation. These challenges have consequently restricted our understanding of the mechanisms linking potential risk factors to the development of DBDs. Innovative causal inference methods and large, representative datasets can (partly) address these limitations and help researchers make inferences about the aetiology of disruptive behaviour.

1.4 Strengthening causal inference

In recent years, there has been an increase in access to large, longitudinal datasets and a theoretical understanding of causal inference and its importance (De Stavola et al., 2022). This has led to many studies incorporating more advanced and nuanced methods to identify causal links between putative risk factors and outcomes. It is vital that research on DBDs also adopts these approaches, given the high prevalence and societal cost of these behaviours (Heeks et al., 2018). Research that can identify the causal, as opposed to correlational, risk factors associated with disruptive behaviour would greatly inform efforts to design more effective interventions to prevent its development.

1.4.1 The counterfactual framework and its assumptions

At the heart of modern causal inference is the counterfactual framework, also known as the potential outcomes framework. The counterfactual framework compares the *(potential) outcomes* of two hypothetical scenarios whereby the same individual is exposed or unexposed to a putative risk factor (Hernán et al., 2004; Rubin, 2007,

2008). Several key technical assumptions are invoked to draw causal inferences within the counterfactual framework. These include but are not limited to (a) no interference, (b) consistency, (c) conditional exchangeability, and (d) positivity. It is assumed that one individual's exposure level does not influence another individual's outcome (*no interference*). It is also assumed that the observed outcome is the same as the potential outcome that would have occurred if the researcher set the exposure level to the same value as the observed exposure level (counterfactual *consistency*). These two assumptions allow us to define the potential outcomes and, from these, the causal contrasts of interest, such as the average treatment effect (ATE), which is the difference in mean potential outcomes (Robins, 1986).

Other assumptions are usually invoked to estimate these causal contrasts. One such assumption is that the observed exposure is independent of the potential outcomes, given covariates, known as the no unmeasured confounding assumption. This is also known as the *conditional exchangeability* assumption, as exposed and unexposed individuals are assumed to be exchangeable, conditional on covariates. Another assumption that can be used is that there is an instrumental variable for the exposure (Hernán & Robins, 2020; discussed in Section 1.4.2.1.2 below). For certain estimation methods, such as those based on the propensity score, the *positivity* assumption is also invoked, which states that each individual in the population has a non-zero probability of receiving each level of exposure, given the covariates. In other words, every individual has a chance of being exposed, conditional on the covariates.

1.4.2 Causal inference methods for observational data

As described in Section 1.3, in RCTs, individuals are assigned to either being exposed or unexposed *experimentally*. In contrast, causal inference methods use observational data and various approaches to control for the lack of randomisation in order to compare exposed and unexposed individuals. Consequently, causal inference

methods are sometimes called *quasi*-experimental studies (Campbell & Stanley, 1966). However, what constitutes a quasi-experimental study is debated (Reeves et al., 2017). Therefore, I use the more general term *causal inference methods* in this thesis. I define causal inference methods as those that aim to estimate causal effects using observational data either by *(a)* relying on an instrument (e.g. regression discontinuity, Mendelian randomisation, difference-in-difference approaches) or *(b)* confounder-control (e.g. extensions to regression-based methods, propensity score matching). Although I make this distinction, there is a degree of overlap between the categories (see Matthay et al., 2020, for an overview). Below, I review how different causal inference methods can exploit features in observational data to deal with common biases arising in epidemiology.

1.4.2.1 Methods that control for the lack of randomness in the exposure by relying on an instrument

1.4.2.1.1 Regression discontinuity

In the absence of experimental randomisation, methods based on natural experiments (e.g. regression discontinuity) can take advantage of naturally occurring events that allocate individuals to certain exposures that either remove or induce risk. Examples of natural experiments include policy changes, such as alterations in educational or financial policies, or natural disasters, such as famines. These methods have long been used in economics and are increasingly used in epidemiology (Bor et al., 2014; Rutter, 2012). However, they require in-depth knowledge of the natural event and the availability of data covering the period it occurred. Consequently, I do not directly employ these methods in my thesis but existing natural experiments are included in my systematic review and meta-analysis (Chapters 2, 3 and 4).

1.4.2.1.2 Instrumental variable (IV) analyses

In epidemiological research, specific genetic variants (single nucleotide polymorphisms; SNPs) that are known or assumed to be associated with an exposure of interest can be used as instrumental variables (IVs) under certain assumptions. In addition to the assumptions described in Section 1.4.1, IVs must also satisfy the assumptions of *(a) relevance*, *(b) exchangeability* and *(c) exclusion restriction*. Relevance states that the IV must be robustly associated with the exposure of interest. Exchangeability states the IV must not share any common causes with the outcome. Exclusion restriction assumes that the IV affects the outcome *only* through the exposure. As genetic variants are randomly assigned at conception, they are unlikely to be associated with environmental confounders. Therefore, causal estimates of the effect of the exposure on the outcome could be obtained even in the presence of unmeasured confounding. An additional advantage of using IVs is that the temporal ordering of the genetic variants and their associated exposure is known, reducing the risk of reverse causality (Pingault et al., 2018). Analyses that utilise SNPs as IVs are known as MR analyses.

1.4.2.1.3 Family-based designs

Family-based designs use data from genetically related individuals, such as siblings and/or twin pairs, where information on shared genetic and environmental factors between family members is known. Sibling comparisons and discordant twin differences design can approximate the counterfactual framework by using the unexposed sibling (or twin) as a "natural match" for their exposed co-sibling (or co-twin; Pingault et al., 2018). This allows researchers to determine whether the risk factor has a causal role in the development of the outcome while controlling for genetic confounding and some, but often not all, environmental confounding (Frisell, 2020; Jaffee, Strait, et al., 2012; Pingault et al., 2018; Sjölander et al., 2022). Many studies investigating risk factors for disruptive behaviour, and therefore included in

my systematic review and meta-analysis (Chapters 2, 3 and 4), have used family-based designs, which are not immune to bias (Rod et al., 2021; Thapar & Rice, 2021).

1.4.2.2 Methods that control for the lack of randomness in the exposure by confounder-control (or standardisation) methods

Compared to the methods above, which exploit exogenous naturally occurring phenomena, other causal inference methods control for the lack of randomisation by controlling for confounders in the analyses. Examples of such methods include fixed-effects and random-intercept cross-lagged models, which leverage longitudinal data to test whether within-individual changes in the exposure predict changes in outcomes independent of stable individual factors. Other examples include extensions of traditional regression-based methods, which account for time-varying confounding, so-called G-methods (Robins, 1986), and include inverse probability weighting of marginal structural models, g-computation, and g-estimation of structural nested models (Chapter 6).

All these methods can only answer causal questions under the assumptions outlined in Section 1.4.1 (Daniel et al., 2013; Naimi et al., 2016). In this thesis, I examine how different causal inference methods have been and can be used to investigate potential risk factors for disruptive behaviour. I will exploit a few causal inference methods to examine whether they produce results which triangulate with existing studies, particularly those that have used causal inference methods that use different sources of data and rely on a different set of assumptions (De Stavola et al., 2022; Lawlor et al., 2017).

1.5 Aims and objectives

This thesis aims to triangulate evidence of potential causal risk factors for DBDs by evaluating the strength of the evidence reported in existing studies, implementing novel causal inference methods, and using unique data sources. I seek to answer the following questions:

1. How can causal inference methods be used to examine the aetiology of disruptive behaviour?

> **Objective**: To summarise the main causal inference methods relevant to answering causal questions about potential risk factors for disruptive behaviour.

2. What have current causal inference methods told us about risk factors for disruptive behaviour?

> **Objective**: To identify all relevant existing studies that examine risk factors for disruptive behaviour disorders using causal inference methods.

3. How can researchers use current evidence to establish the magnitude of the effects of parenting practices on* DBDs*?

> **Objective**: To combine and summarise all relevant evidence on the relationship between parenting practices and disruptive behaviour by meta-analysing studies that use causal inference methods.

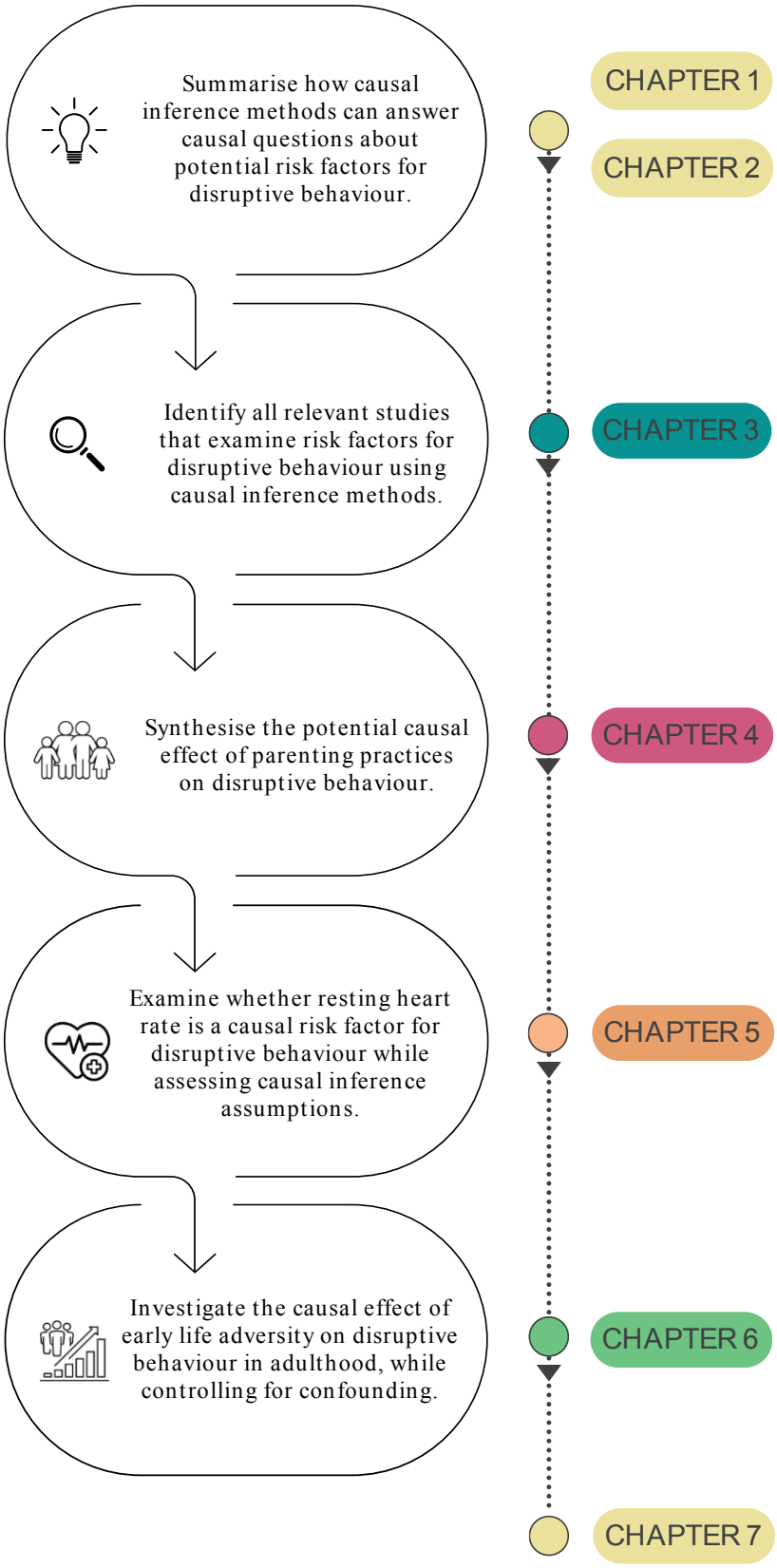
4. How can instrumental variable analyses be used to evaluate the influence of biological risk factors on* DBDs*?

> **Objective**: To exploit Mendelian randomisation, a genetically informed causal inference method, to examine a biological risk factor for disruptive behaviour, resting heart rate, while assessing the feasibility of the assumptions needed for causal inference.

5. How can population-based cohorts and administrative data best be used to address causal questions on the aetiology of* DBDs*?

Objective: To investigate the causal effect of early life adversities on disruptive behaviour in adulthood using routinely collected data and controlling for confounding using g-methods.

1.6 Thesis structure



In Chapter 2, I describe a study protocol for a systematic review that aims to identify all risk factors for DBDs from the current causal inference literature. I will also define the main causal inference methods relevant to answering causal questions about potential risk factors for disruptive behaviour and outline the steps necessary to identify all relevant existing studies.

In Chapter 3, I summarise the findings from my systematic review of risk factors for DBDs using causal inference methods. I also outline how the results from my systematic review informed my decision on which risk factors and methods to focus on in this thesis.

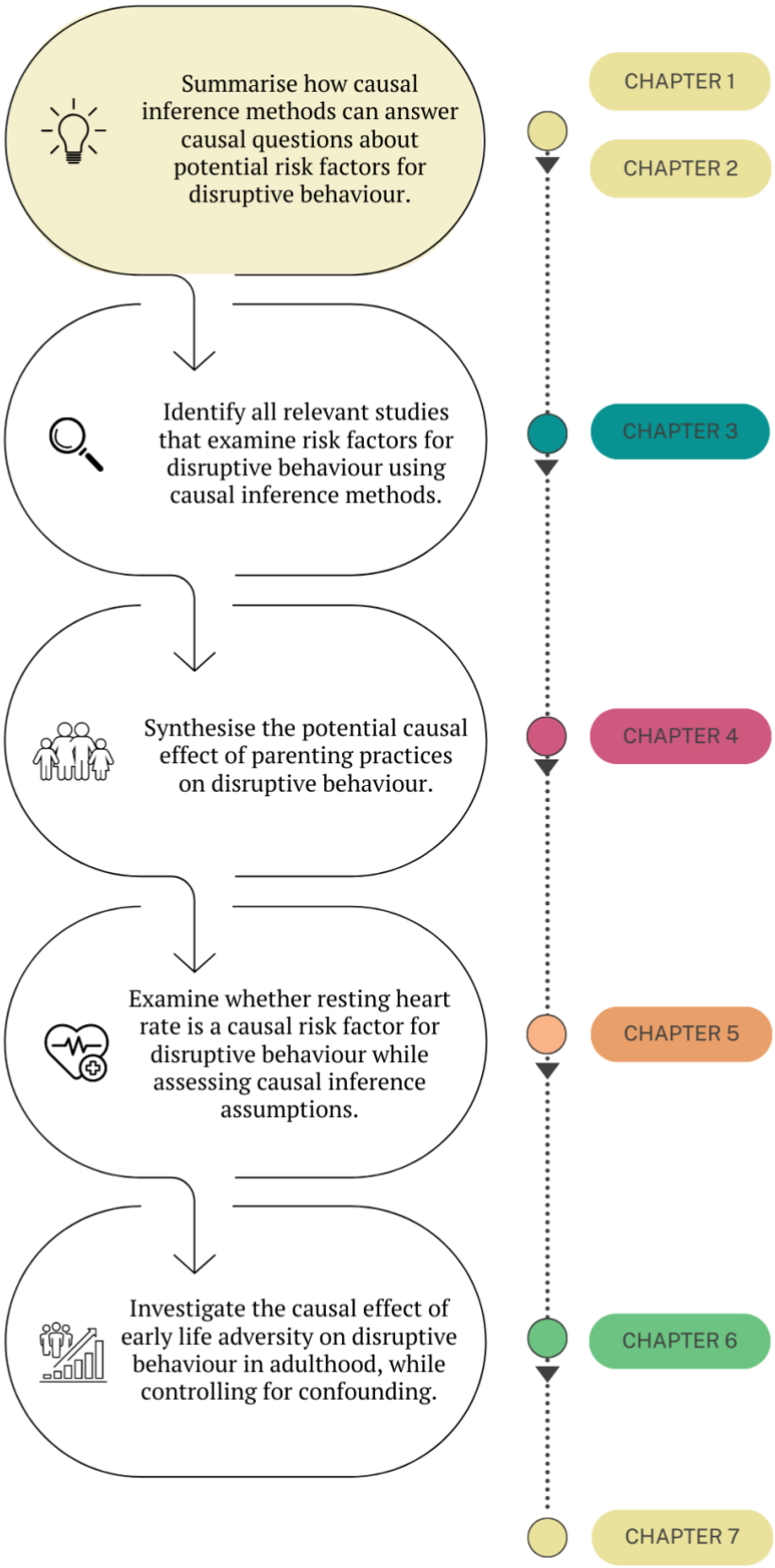
In Chapter 4, I exploit the wealth of existing studies that use causal inference methods to investigate the effect of parenting practices on disruptive behaviour. I present an example of how this evidence can be synthesised in a meta-analysis to estimate a pooled causal effect and examine potential sources of bias.

In Chapter 5, I use publicly available genetic data to examine a biological risk factor for disruptive behaviour, resting heart rate, using Mendelian randomisation analyses. I will also assess the feasibility of the assumptions needed for causal inference using instrumental variable approaches.

In Chapter 6, I analyse routinely collected registry-based data from a whole-population cohort to identify common clusters of early life adversity using group-based multiple-trajectory modelling. I will then investigate whether exposure to different clusters of early life adversities influences the risk of DBD symptoms throughout adolescence and adulthood.

In Chapter 7, I summarise the key learnings from this thesis and discuss how the empirical Chapters have addressed the questions posed in this thesis. I will also discuss potential next steps.

THESIS STRUCTURE



2 PROTOCOL FOR A SYSTEMATIC REVIEW AND META-ANALYSIS OF STUDIES USING CAUSAL INFERENCE METHODS TO INVESTIGATE RISK FACTORS FOR DISRUPTIVE BEHAVIOURS

Chapter overview

In Chapter 1, I outlined the current evidence on risk factors for DBDs and described common threats to the validity of these findings. In this Chapter, I describe a protocol for systematically identifying, evaluating, and, where appropriate, synthesising evidence from existing causal inference studies on risk factors for DBDs. I will also define the key terms used in this thesis.

Publication status: This Chapter is based on an article published in the *BMJ Open*: Karwatowska, L., Russell, S., Solmi, F., Stavola, B. L. D., Jaffee, S., Pingault, J.-B., & Viding, E. (2020). Risk factors for disruptive behaviours: Protocol for a systematic review and meta-analysis of quasi-experimental evidence. BMJ Open, 10(9), e038258. https://doi.org/10.1136/bmjopen-2020-038258

2.1 Background

As discussed in Chapter 1, individuals who display disruptive behaviours in childhood (including conduct problems [CP], conduct disorder [CD] and oppositional defiant disorder [ODD]) engage in a range of repetitive and troublesome behaviours, such as lying, fighting, and stealing. Due to the emotional, social, and economic burden these behaviours place on society (Heeks et al., 2018; Rivenbark et al., 2018), there has been a long-standing interest in understanding what factors increase (or decrease) risk. However, much of the existing evidence draws upon research that has used

unrepresentative samples (e.g., high-risk or clinical samples), has explored a restricted number of risk factors and/or has identified (non-causal) associations between putative risk factors and outcomes due to methodological and statistical limitations (Fairchild et al., 2019; Rutter, 2007). In recent years, causal inference methods have gained popularity as they often use large, representative samples and, under certain assumptions, can estimate causal effects between risk factors for which classical randomised control trials (RCTs) are either unethical, impractical or too costly (Bärnighausen, Tugwell, et al., 2017).

Triangulation is the process of combining evidence from different sources to answer causal questions more reliably (De Stavola et al., 2022; Lawlor et al., 2017). Although triangulation is a relatively new term, it is not a new concept in epidemiology. Indeed, there is a long history of epidemiologists effectively triangulating evidence by conducting systematic reviews and meta-analyses. However, a central concept of triangulation is that the studies have different and unrelated sources of bias. Therefore, evidence from meta-analyses which combine results from similarly biased observational evidence will inevitably also be biased. Instead, in a seminal paper, Jaffee and colleagues published a narrative review on the evidence of risk factors for antisocial behaviour, a common symptom of DBDs, from studies which used causal inference methods (Jaffee, Strait, et al., 2012). The authors concluded that there was evidence that harsh parental discipline, maltreatment, parental divorce, adolescent motherhood, maternal depression, parental antisocial behaviour, peer deviance and poverty all had causal effects on antisocial behaviour. On the other hand, there was no evidence of causal effects for smoking during pregnancy, paternal depression, parental alcohol use or neighbourhood disadvantage.

However, the review only considered a limited number of risk factors and causal inference methods and only synthesised the evidence qualitatively instead of

quantitatively. Many risk factors proposed to have causal effects on disruptive behaviour (Chapter 1, Section 1.2) were not included in the Jaffee review, including other perinatal factors, psychophysiological factors, and peer problems. In addition, the studies included were mainly genetically informed family-based methods and did not include other causal inference methods, such as regression discontinuity, interrupted time series, or instrumental variable (IV) analyses. In recent years, the number of studies using causal inference methods to examine risk factors for DBDs has increased considerably. Therefore, another review of the evidence is timely and important.

More than ten years on and, to my knowledge, only two other studies have synthesised evidence of risk factors for DBDs from causal inference methods. One review included studies that controlled for genetic effects while examining the relationship between prenatal smoking, alcohol and caffeine exposure and externalising problems, including CD and ODD (Haan et al., 2022). The authors did not identify enough high-quality studies to conduct a meta-analysis. After systematically reviewing the evidence, they concluded that there was *some* evidence of an association between prenatal smoking and CD but not enough evidence for prenatal smoking and ODD or alcohol and caffeine exposure with either CD or ODD. Another recent review, which I co-authored, looked at the effect of childhood maltreatment on mental health problems, including externalising (*k* [number of studies] = 2) and CP (*k* = 7; Baldwin et al., 2023). The pooled meta-analytic estimate was much smaller than those previously reported in meta-analyses of studies that did not use causal inference methods but suggested that childhood maltreatment has a small causal effect on externalising behaviours (Cohen's d = 0.35; *95% CI* = 0.24, 0.46). Both previous meta-analyses focussed on specific risk factors for DBDs. They also used a broad definition of externalising (EXT) behaviour, for example, including

attention deficit hyperactivity disorder (ADHD; Haan et al., 2022) and alcohol and drug abuse (Baldwin et al., 2023).

In this Chapter, I outline a protocol for a systematic review and meta-analysis published in 2020 to identify all existing studies that use causal inference methods to investigate putative risk factors for DBDs. I considered an inclusive range of outcomes for disruptive behaviour, including diagnostic (e.g., CD, ODD, antisocial personality disorder [ASPD] and dissocial personality disorder [DPD]) and continuous measures that confer sub-clinical symptoms (e.g., CP). To limit the review, I did not include general symptoms (e.g., antisocial behaviour, delinquency, crime) which are not specific to DBDs but also shared with other disorders. Risk factors were not selected *a priori,* which allowed any putative risk factor for DBDs to be included. I broadly defined causal inference methods as those that use observational data, in contrast to RCTs which use experimental data, to infer causality either in their study design, using an instrument, or in their analyses, using confounder-control methods. Given the increased awareness and use of causal inference methods in recent years, I hypothesised that there would be sufficient studies to perform separate meta-analyses on most identified risk factors.

By systematically combining and summarising all relevant literature, the current review aimed to:

- Identify risk factors for DBDs from causal inference methods and examine whether these results indicate evidence of causal effects.
- If so, establish whether the results vary by participant and study characteristics, specifically:
    - Participant sex and age.
    - Type of disruptive behaviour outcome.

- Causal inference method.
- Study quality.
- Any other characteristics relevant to the specific risk factor of interest.

2.2 Methods

The protocol for my systematic review was preregistered with the PROSPERO database [CRD42020169313] and published in the BMJ Open (Karwatowska et al., 2020) using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) guidelines (Shamseer et al., 2015; Appendix A).

2.2.1 Definition of key terms

2.2.1.1 Causal inference methods

As discussed in Chapter 1, the definition of causal inference methods used in this thesis are those methods that aim to estimate population-level causal effects using observational data either by *(a)* relying on an instrument (e.g. regression discontinuity, Mendelian randomisation, difference-in-difference approaches) or *(b)* confounder-control methods (e.g. extensions to regression-based methods, propensity score matching). In the tables below, I define the key terms for study designs (Table **2.1**), analyses (Table **2.2**) and features (Table **2.3**) associated with causal inference methods. These definitions were informed and adapted from the causal inference literature (Bärnighausen, Tugwell, et al., 2017; Bor et al., 2014; Davies et al., 2017; Funk et al., 2011; Gunasekara et al., 2014; Hernán, 2004; Jandoc et al., 2015; Pingault et al., 2018; Rubin, 2007, 2008; Rutter, 2012) and were used in my database searches (Section 2.2.3 below).

Table 2.1 *Definitions of key designs commonly used in causal inference methods.*

Term	Definition
Adoption study	Adoption separates genetically related parents and children and places children in a different rearing environment. Adoption studies compare associations between exposures and outcomes in parents and their children that are genetically related (not adopted) and genetically unrelated (adopted) to them. (Pingault et al., 2018).
In vitro fertilisation study	In vitro fertilisation can use parental gametes (genetically related) or donor gametes (genetically unrelated) for fertilisation. Similar to adoption studies, in vitro fertilisation studies compare associations between exposures and outcomes in parents and children that are genetically related and genetically unrelated. (Pingault et al., 2018).
Genetically informed methods	Study designs or statistical analyses that use genetic information (e.g. known genetic relationships between twins) or data on genetic variations. (Pingault et al., 2018).
Natural experiment	Natural experiments use randomly occurring circumstances (e.g. lottery win, policy, or law change) as an exposure, which is assigned "as random"(Rutter, 2012). True natural experiments are very unusual. Therefore they are usually described as "quasi-natural experiments" or observational studies with exogenous exposures.
Sibling study	On average, siblings share 50% of their segregated genetic material. Similar to twin studies, some sibling studies compare outcomes in exposed versus non-exposed siblings. (Pingault et al., 2018).
Twin study	On average, dizygotic twin pairs share 50% of their segregated genes compared to monozygotic twin pairs which share 100% of their genetic material. A twin that is non-exposed to a risk factor represents a natural match to their exposed co-twin. Therefore, some twin studies compare outcomes in exposed versus non-exposed pair members. (Pingault et al., 2018).

Table 2.2 *Definitions of key types of analyses commonly used in causal inference methods.*

Term	Definition
Difference in difference study / controlled before and after study	Difference in difference designs, also known as controlled before and after studies, are a type of fixed effect study. The difference before and after the exposure in the exposed group is compared to the same period of time in the non-exposed group. The exposure effect is the difference between these differences. If an exposure has a harmful effect, the outcome will occur more rapidly in individuals who receive the exposure than in individuals who do not. (Bärnighausen, Tugwell, et al., 2017).
Fixed effects	Applied to longitudinal data with repeated measures, fixed effects methods model within-individual changes over time (i.e. variation in an individual's exposures and outcomes), as opposed to between-individual changes (i.e. variation across individuals), to remove time-invariant confounding, with each individual acting as their own control. Difference-in-difference, controlled before-and-after, experience sample and ecological momentary assessment are all examples of fixed effect analyses. (Gunasekara et al., 2014).
Instrumental variable analysis	Analyses that use variables that are associated with an exposure of interest (*relevance*), do not share any common causes with the outcome (*exchangeability*) and affect the outcome only through the exposure (*exclusion restriction*), also known as instrumental variables. These variables can be any traits that meet the three instrumental variable assumptions, for example genetic variants (e.g., Mendelian randomisation; Davies et al., 2017).
Interrupted time series analysis	Interrupted time series methods use observational data collected over equally spaced intervals before and after an intervention, that is exogenous to the time series, e.g. a "natural experiment" in the real world setting. The effect of the intervention is evaluated by examining whether the data pattern (e.g. the level and slope) observed post-intervention is different from that observed pre-intervention. (Jandoc et al., 2015).
Regression discontinuity analysis	If treatment allocation is based on whether a patient scores below or above a predetermined cut-off value, as opposed to randomisation, then the intervention will be randomly assigned for patients close to the threshold. (Bor et al., 2014). See also "Sharp/fuzzy design" below.

Table 2.3 *Definitions of key features commonly used in causal inference methods.*

Term	Definition
Causal effect	An exposure has a causal effect on the outcome if the outcome differs when the exposure is present compared to when the exposure is absent, all other things being equal. (Hernán, 2004).
Counterfactual framework	The comparison of hypothetical scenarios whereby the same individual is either exposed or unexposed to a risk factor. Also known as the potential outcomes framework. (Hernán, 2004; Rubin, 2007, 2008).
Doubly robust estimation	Doubly robust estimation combines two models: outcome regression and propensity score modelling. Individually, these two methods lead to unbiased estimators of the causal effect only if the respective model is correctly specified; when combined, through doubly robust estimation, only one of the two models needs to be correctly specified to obtain an unbiased effect estimator. (Funk et al., 2011).
Heckit model/Heckman sample selection	Similar to selection/selectivity models, these handle non-ignorable missing data. Heckit/Heckman selection models assume *(a)* a joint distribution for the missingness and outcome processes and *(b)* validity in the instrument. If these assumptions are met, these models can correct bias from non-randomly selected samples.
Matching study	Researchers can attempt to create a reasonable counterfactual by accounting for confounders via matching exposed and non-exposed participants on key variables. Propensity score matching approaches can select appropriate matches (either to cases, non-cases or both), leading to different causal effects. Quasi-experimental designs are often combined with propensity score matching approaches. (Rutter, 2012). See "Propensity score."
Potential outcome	The outcome would occur had the exposure been set to a particular value. See "Counterfactual framework."
Propensity score	A propensity score is the probability of being exposed conditional on the confounders. It has the advantage of reducing, a potentially large, number of confounders into a scalar that contains all information that is relevant for the exposure assignment in relation to the outcome. The propensity score is used as an additional covariate in outcome regression, or as a stratifying or matching variable. Inverse probability weights derived from propensity scores also remove confounding by recreating a

	pseudo-sample where there is no confounding. See "Matching study."
Selection/selectivity model	A model that deals with samples that are non-randomly selected, and therefore non-representative of the target population. For example, studies affected by non-ignorable missing data.
Sharp design/Fuzzy design	These are features of regression discontinuity designs. A sharp discontinuity regression design exploits exogenous changes to the value of an exposure /intervention to estimate its causal effect on an outcome. These changes are usually triggered by overtaking a particular (sharp) threshold in a continuous endogenous variable. Since the comparison with the threshold may be affected by random error, individuals with values near the threshold can be viewed as being "as good as" randomly allocated to the exposure and analysed as if they were in an RCT. In a fuzzy regression discontinuity design, the threshold does not need to be defined as a sharp discontinuity if the probability of exposure/intervention assignment differs among those near the threshold.

2.2.1.2 Disruptive behaviour disorders

The literature uses multiple terms for disruptive behaviour interchangeably. For the purposes of the current thesis, I define disruptive behaviour disorders (DBDs) as either CD, CP, or ODD in childhood and ASPD, DPD, and psychopathy in adulthood. I describe these terms further in Table **2.4**.

Table 2.4 *Definitions of key terms used to describe disruptive behaviour disorders and related symptoms.*

Term	Definition
Disruptive behaviour disorders	A range of repetitive and troublesome behaviours, such as lying, fighting, and stealing. It is sometimes used interchangeably with "externalising disorders."
Antisocial personality disorder (DSM) / dissocial personality disorder (ICD)	A diagnosis which involves a life-long pattern of antisocial behaviour as well as irritability and remorselessness. By definition, a diagnosis of ASPD involves exhibiting conduct disorder in childhood.
Conduct disorder	A formal diagnosis whereby an individual displays repetitive and persistent patterns of antisocial, aggressive, or defiant behaviour that amounts to significant and persistent violations of age-appropriate social expectations (National Institute for Health and Care Excellence, 2013) that are diagnosable as defined by the DSM-5.
Conduct problems	Disruptive behaviour that does not necessarily meet the threshold for diagnosis of conduct disorder.
Oppositional defiant disorder	A formal diagnosis whereby an individual exhibits defiant and disobedient behaviour towards others as opposed to conduct disorder, whereby behaviours violate the rights of others and/or societal expectations.
Psychopathy	Psychopathy is characterised by high levels of antisocial behaviour, low levels of anxiety and high levels of attention seeking. It is a specifier for antisocial personality disorder (ASPD) in the DSM-5.

2.2.2 Eligibility criteria

Studies meeting *all* the following criteria were included in the review:

- The study only included human participants, i.e., did not include non-human animals.
- To ensure consistency in the definitions, the study must have included at least one clearly defined measure of a risk factor and at least one clearly defined measure of disruptive behaviour. The risk factor must have occurred before the outcome.
- Effect sizes must have been reported, or there must have been enough numerical information to calculate them. I contacted study authors to request additional data for studies that did not meet this criterion.
- The study could have been conducted in any country, but it must have been published in English for practical reasons.
- The study must have been published after 1980 to maintain consistency in the definition of disruptive behaviours.
- The study must have used a causal inference method (defined in Tables 2.1-2.3).

Studies meeting *any* of the following criteria were excluded from the review:

- The study did not meet the above inclusion criteria.
- The study was a case report, clinical trial, editorial, letter to the editor, systematic review, or meta-analysis.
- The study used populations selected based on participant physical health problems, such as cancer, seizures, surgery, low gestational age, etc., which were not the focus of the current review.
- The study used populations selected on other diagnosed developmental disorders (e.g., language disorders, learning disorders, motor disorders, autism spectrum disorders, etc.) or mental health diagnoses (e.g., schizophrenia,

depression, bipolar, etc.). Disruptive behaviour can share symptomology with other developmental and mental health disorders, but these go beyond the scope of the present review.

2.2.3 Search strategy

An electronic search was conducted to identify all relevant studies from the 1st of January 1980 until the 1st of January 2021. The electronic databases were selected either because the database was relevant for the current review's research question, e.g., PsycINFO, and/or because the database is frequently used in literature searches, e.g., MEDLINE. I also included the Web of Science Core Collection database as eight articles in the review conducted by Jaffee and colleagues (2012) were not available in the Ovid databases but were available in Web of Science. As such, I conducted systematic searches of the following databases:

- Ovid
- MEDLINE In-Process & Other Non-Indexed Citations and Daily
- EMBASE
- PsycINFO
- Web of Science Core Collection

The search terms used for the causal inference methods were adapted from a paper by Glanville and colleagues (2017) to include genetically informed causal inference methods, such as twin designs and Mendelian randomisation. The search terms for disruptive behaviours were selected to include diagnostic terms. However, they did not include terms for symptoms associated with disruptive behaviour, such as antisocial behaviour and delinquency, as these are not specific to DBDs. Appendix A provides additional information on the keywords (Table 1) and database search techniques (Table 2).

2.2.4 Study selection

Citations were imported into EPPI-Reviewer 4 (EPPI-4; Thomas et al., 2010), a data management software. EPPI-4 includes a machine learning process which aims to reduce the time taken to screen titles and abstracts by prioritising unscreened articles based on the reviewers' previous screening decisions. Specifically, the EPPI-4 software assesses the frequency of words in the inclusion compared to the exclusion categories. The consistency between reviewers' screening decisions is checked periodically, and the list of unscreened references was refreshed, which allowed the machine learning software to prioritise unscreened items based on relevance denoted from inclusion and exclusion codes. Two independent reviewers (L.K. and F.S.) completed the initial screening of the abstracts and titles using EPPI-4. Any references categorised as included by both reviewers after screening on title and abstracts had their full texts screened for inclusion by two independent reviewers (LK and FS). Any uncertainties over the inclusion/exclusion of studies were resolved by team consultation.

2.2.5 Data extraction

Multiple independent reviewers (L.K. and F.S., B.L.D.S. or J.R.B.) conducted data extraction and quality assessment on any references included after the full-text screening. Any discrepancies were resolved through discussion, and missing data or incomplete information was requested from the study authors. The following information was extracted from the studies using a data extraction form: study reference, project name and country, study design (e.g. cohort study, etc), participant information (e.g. number, ethnicity, age at measurement, etc.), main exposure and outcome measurement features (e.g. measurement tool, rater, age at measurement, etc.), confounders, additional risk factors, additional outcomes, average effect size and other relevant quantities (e.g. estimate, standard error, sample size, exclusions, attrition, etc.).

2.2.6 Quality assessment

I adapted the Newcastle-Ottawa scale (Wells et al., 2000) to include questions relevant to causal inference methods (Appendix A, Table 3). Additional/adapted questions included control for environmental and genetic confounders (Q5 and Q6), whether the exposure and outcome were reported by different informants (Q8), and whether the exposure and outcome were assessed longitudinally (Q9).

An overall score was derived by summing the scores across all items (highest score = 10), and the 33rd and 66th percentiles were used to categorise the studies into one of three categories: very high-risk (score below 5.5), high-risk (score between 5.5 and 7) or high-quality (score above 7). For studies that reported multiple effect estimates in different categories (e.g., high-quality and high-risk), we gave the study an overall rating corresponding to the highest category (e.g., high-quality). One author (L.K.) coded the study quality and discussed any questions with two team members (B.L.D.S. and J-B.P.).

2.2.7 Strategy for data synthesis

Depending on the amount and quality of information provided in the included studies, I conducted either qualitative syntheses (i.e., a systematic review; Chapter 3) or quantitative syntheses (i.e., a meta-analysis; Chapter 4). Meta-analyses were only deemed appropriate if a minimum of three studies reported effect estimates on a particular risk factor and a particular outcome that were sufficiently homogenous to lead to sensible summary estimates.

I conducted multi-level linear random-effects models (Assink & Wibbelink, 2016; described in more in detail in Chapter 4, Section 4.2.6) to account for study heterogeneities. The resulting pooled estimates were reported together with measures of their dispersion. The I^2 statistic was used to quantify heterogeneity, with

an I^2 of more than 50% indicating moderate heterogeneity (Higgins & Thompson, 2002). The Metafor package in R (Viechtbauer, 2010) was used to conduct the analyses.

2.2.8 Analysis of subgroups

When sufficient data were available, meta-analyses were used to assess the specificity of pooled effects by examining whether the effects varied across pre-specified "subgroups." For the current review, subgroups were defined according to a variety of participant (e.g. outcome, sex, and age of onset) and study characteristics (e.g. study design, analytical method, in particular the level of confounder adjustment, and data quality). The direct examination of heterogeneities across subgroups was decided depending on the information provided by the studies included in the meta-analysis.

Key points

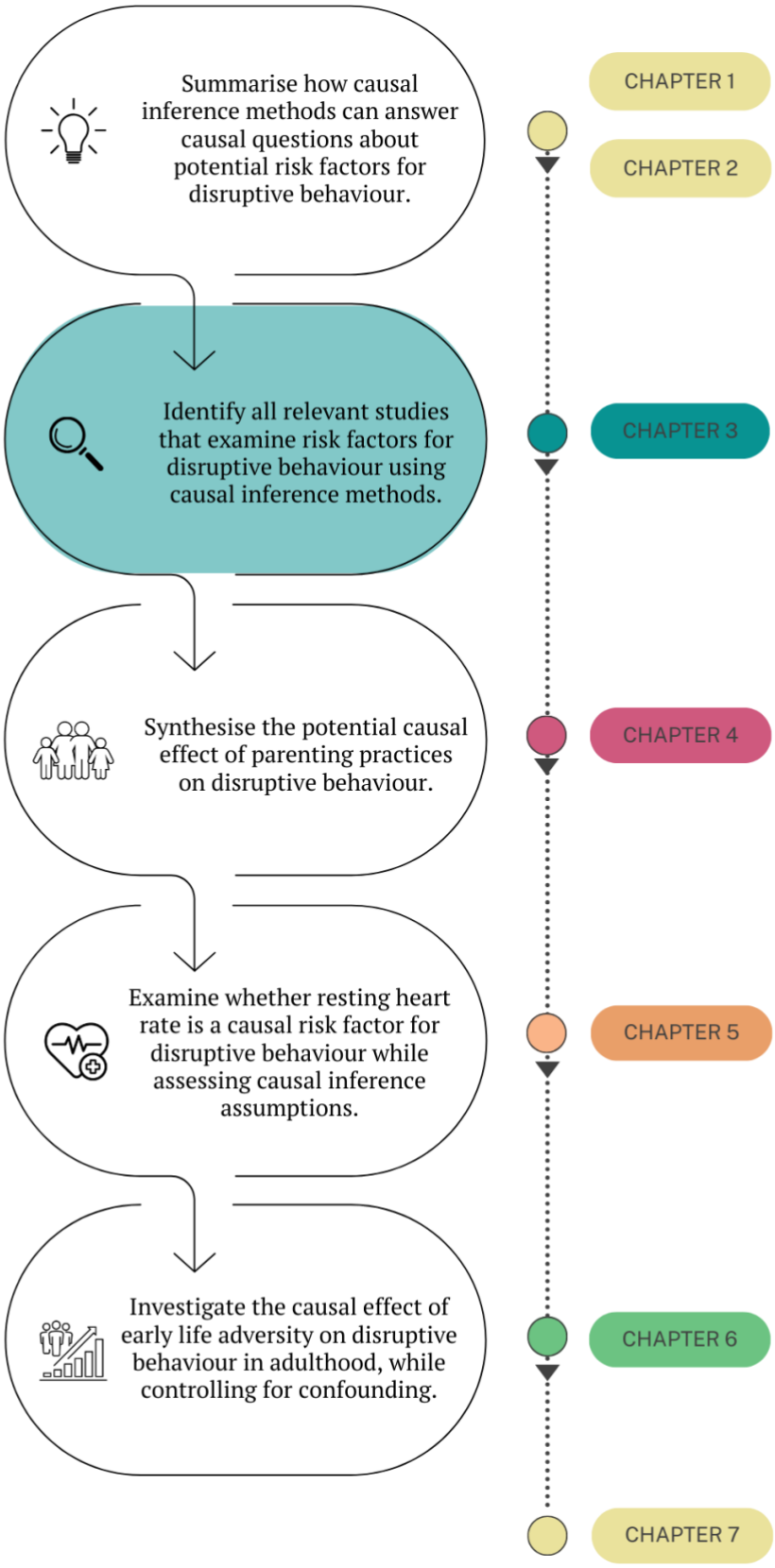
1. In this Chapter, I outlined a study protocol of the first systematic review and meta-analysis to evaluate potential causal effects between all risk factors for disruptive behaviours.

2. I define causal inference methods as those that estimate population-level causal effects using observational data either by *(a)* relying on an instrument (e.g. regression discontinuity, Mendelian randomisation, difference-in-difference approaches) or *(b)* confounder-control methods (e.g. extensions to regression-based methods, propensity score matching)

3. I define DBDs in childhood as conduct disorder, conduct problems, oppositional defiant disorder and in adulthood as antisocial personality disorder, dissocial personality disorder and psychopathy.

4. The results from this review will identify the most probable causal risk factors for disruptive behaviours and highlight potential candidates for future research.

THESIS STRUCTURE



3 SYSTEMATIC REVIEW OF STUDIES USING CAUSAL INFERENCE METHODS TO INVESTIGATE RISK FACTORS FOR DISRUPTIVE BEHAVIOURS

Chapter overview

In Chapter 2, I outlined a study protocol to identify and synthesise existing evidence of risk factors for DBDs from the causal inference literature. In this Chapter, I will summarise the results from the database searches, provide an overview of the studies deemed eligible after full text screening and explain the rationale behind the risk factors considered in this thesis.

Publication status: This Chapter is in preparation for submission to *Psychological Bulletin*.

3.1 Background

The number of studies using causal inference methods has increased considerably since the publication of a seminal narrative review of studies using causal inference methods to investigate risk factors for antisocial behaviour, a common symptom of disruptive behaviour disorders (DBDs). As such, a quantitative synthesis of the results from causal inference methods, which rely on different information sources and assumptions, is timely and important to triangulate evidence on which factors cause DBDs.

3.2 Methods

A detailed description of the methods used in this review, including eligibility criteria and definitions of key terms, is available in Chapter 2. I searched Embase, PsycINFO, MEDLINE and Web of Science Core Collection for peer-reviewed studies written in English and published from the 1st of January 1980 to the 1st of January 2021.

3.3 Results

3.3.1 Search results

As shown in Figure **3.1**, 5298 articles were identified after removing duplicates, 627 of which were included in the full-text screening. Of these, 460 were excluded due to not using a causal inference method (*k* [number of studies] = 406), ineligible outcome measure (*k* = 30) or ineligible risk factor measure (*k* = 14). In total, 167 studies were included in the current review. Further details on each study can be found in Appendix B.

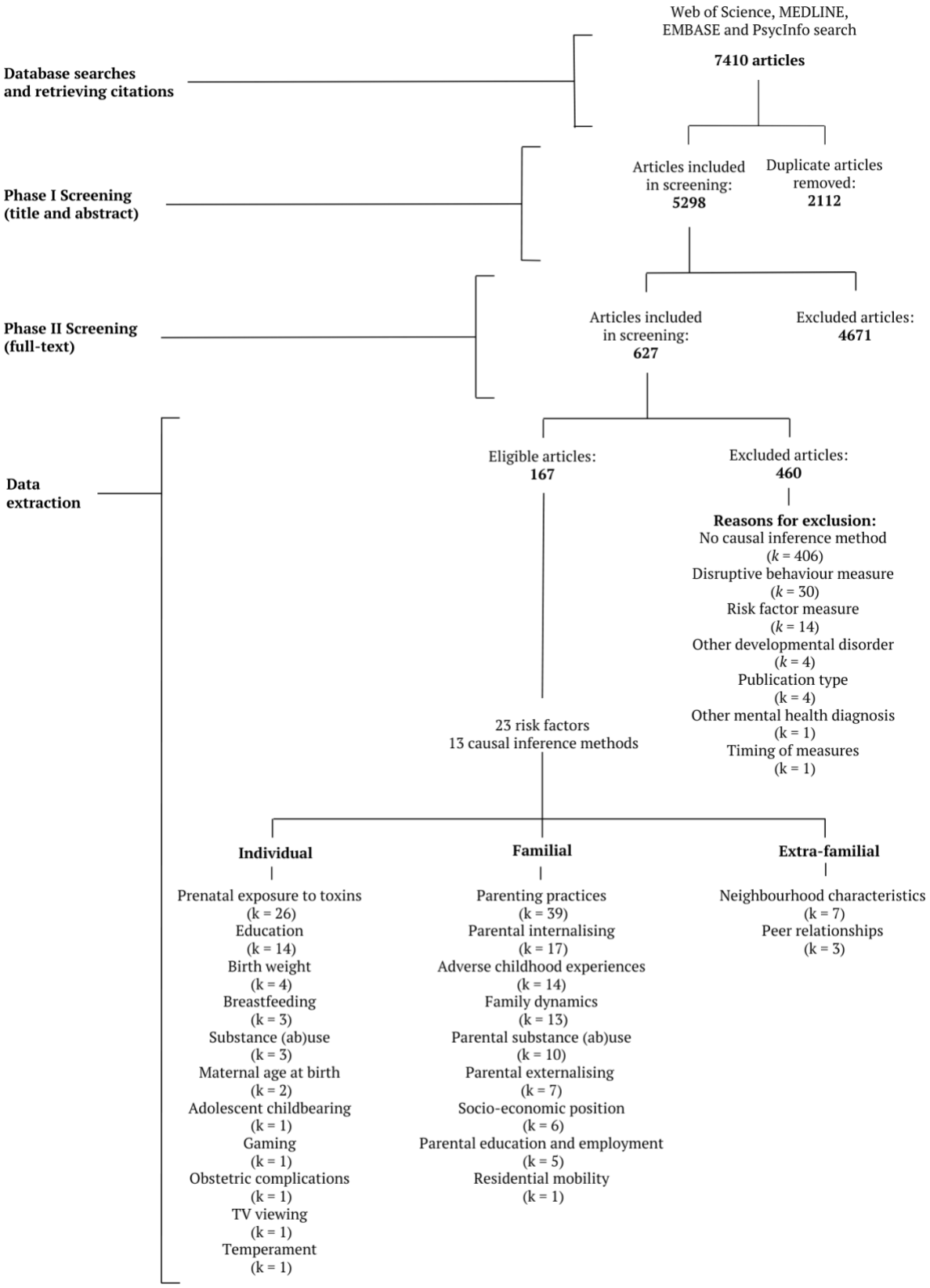


Figure 3.1 *PRISMA flow diagram of the studies included in the systematic review of all risk factors for DBDs.*

3.3.2 Descriptive analyses

The studies included 934,876 unique individuals from 73 distinct cohorts across 18 countries. Most studies used data from cohorts in the United States of America (USA; *k* = 91), the United Kingdom (UK; *k* = 26) and Norway (*k* = 14). The included studies examined 23 putative risk factors for disruptive behaviour using 13 different causal inference methods. A heat map of these studies indicates that the majority focussed on three risk factors (parenting practices [*k* = 39], prenatal exposure to toxins [*k* = 26] and parental internalising symptoms [*k* = 17]) and three causal inference methods (propensity score matching analyses [*k* = 35], discordant sibling study design [*k* = 32] and adoption study [*k* = 29]; Figure **3.2**).

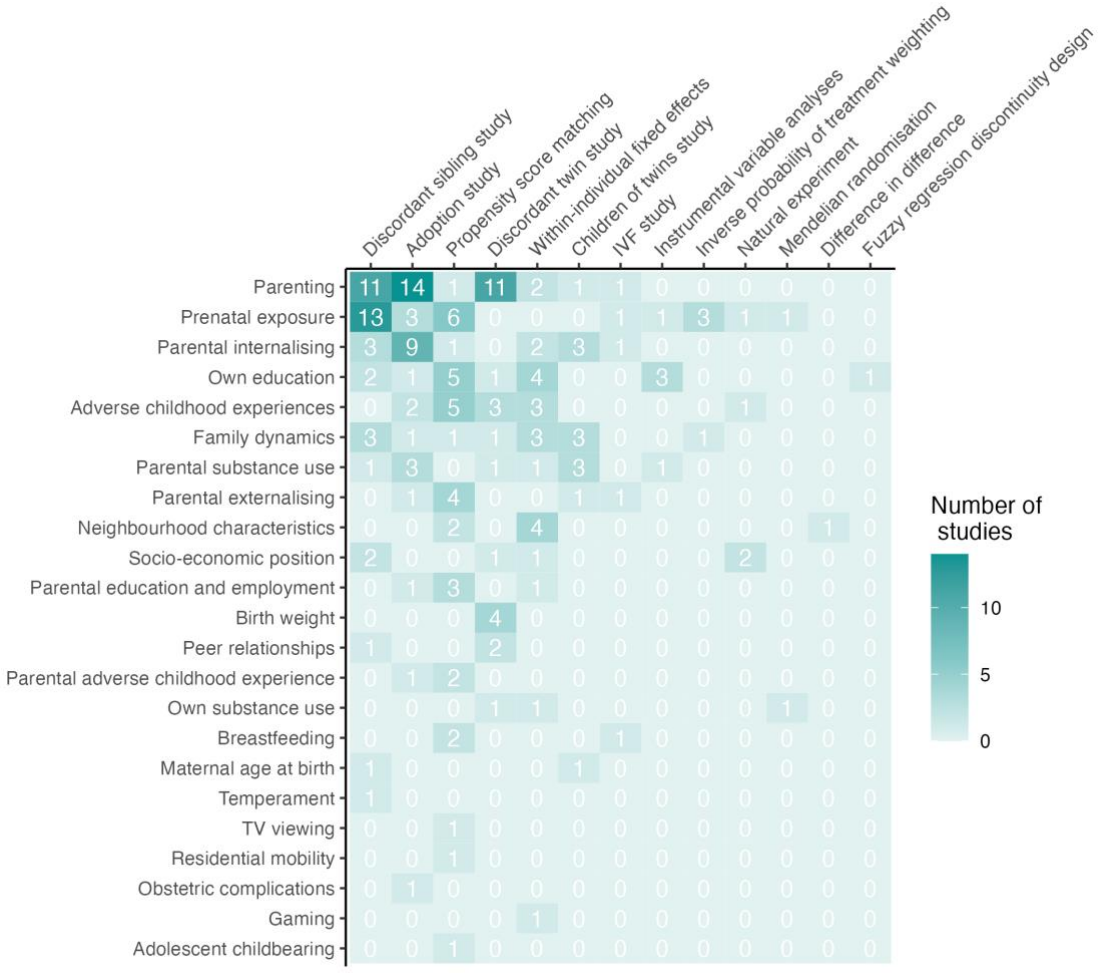


Figure 3.2 *Heat map of the risk factors identified in the current systematic review and the types of causal inference methods used to examine their influence on disruptive behaviour disorder symptoms.*

Regarding individual-, familial- and extra-familial factors (outlined in Chapter 1, Section 1.2), 59 studies focused on individual-level factors, 105 on familial-level factors and 10 on extra-familial-level factors. Further information is reported in Table **3.1**.

Table 3.1 *Descriptive summary of the studies included in the systematic review.*

Characteristic	***k* (%)**
Risk factor	
Parenting practices	39 (21.4%)
Prenatal exposure to toxins	26 (14.3%)
Parental internalising symptoms	17 (9.3%)
Adverse childhood experiences	14 (7.7%)
Education	14 (7.7%)
Family dynamics	13 (7.1%)
Parental substance use	10 (5.5%)
Neighbourhood characteristics	7 (3.8%)
Parental externalising symptoms	7 (3.8%)
Socioeconomic position	6 (3.3%)
Parental education and employment	5 (2.7%)
Birth weight	4 (2.2%)
Breastfeeding	3 (1.6%)
Substance use	3 (1.6%)
Parental adverse childhood experience	3 (1.6%)
Peer relationships	3 (1.6%)
Maternal age at birth	2 (1.1%)
Adolescent childbearing	1 (0.5%)
Gaming	1 (0.5%)
Obstetric complications	1 (0.5%)
Residential mobility	1 (0.5%)
TV viewing	1 (0.5%)
Temperament	1 (0.5%)
Causal inference method	
Propensity score matching	35 (20.6%)
Discordant sibling study	32 (18.8%)
Adoption study	29 (17.1%)
Discordant twin study	24 (14.1%)
Within-individual fixed effects	16 (9.4%)

Characteristic	*k* (%)
Children of twins study	12 (7.1%)
IVF study	5 (2.9%)
Instrumental variable analyses	5 (2.9%)
Inverse probability of treatment weighting	4 (2.4%)
Natural experiment	4 (2.4%)
Mendelian randomisation	2 (1.2%)
Difference in difference	1 (0.6%)
Fuzzy regression discontinuity design	1 (0.6%)
Country	
USA	91 (52.6%)
UK	26 (15%)
Norway	14 (8.1%)
Australia	11 (6.4%)
Canada	4 (2.3%)
Sweden	4 (2.3%)
Ireland	3 (1.7%)
China	2 (1.2%)
Denmark	2 (1.2%)
France	2 (1.2%)
Japan	2 (1.2%)
Switzerland	2 (1.2%)
Chile	1 (0.6%)
Finland	1 (0.6%)
Israel	1 (0.6%)
New Zealand	1 (0.6%)
Puerto Rico	1 (0.6%)
The Netherlands	1 (0.6%)
Not reported	4 (2.3%)
Cohort	
Early Growth and Development Study (EGDS)	14 (8.1%)
Norwegian Mother, Father, and Child Cohort Study (MoBa)	12 (7.0%)
Fragile Families and Child Wellbeing Study (FFCWS)	10 (5.8%)
National Longitudinal Study of Youth (NLSY79)	9 (5.2%)
Twins Early Development Study (TEDS)	8 (4.7%)
Sibling Interaction and Behavior Study (SIBS)	7 (4.1%)
Australian Twin Register (ATR)	5 (2.9%)
Cardiff IVF study (C-IVF)	4 (2.3%)
Early Childhood Longitudinal Study–- Kindergarten Cohort (ECLS-K)	4 (2.3%)
Minnesota Twin Family Study (MTFS)	4 (2.3%)

Characteristic	*k* (%)
Australian twin-family study of alcohol use disorder (OZALC)	4 (2.3%)
Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)	3 (1.7%)
Growing Up in Ireland Child Cohort (GUI)	3 (1.7%)
Millennium Cohort Study (MCS)	3 (1.7%)
Vietnam Era Twin (VET)	3 (1.7%)
Beijing Twin Study (BeTwiSt)	2 (1.2%)
Collaborative Perinatal Project (CPP)	2 (1.2%)
Environmental Risk (E-Risk)	2 (1.2%)
Family Life Project (FLP)	2 (1.2%)
Great Smoky Mountains Study (GSMS)	2 (1.2%)
Northeast-Northwest Collaborative Adoption Projects (N2CAP)	2 (1.2%)
National Institute of Child Health and Human Development (NICHD)	2 (1.2%)
National Longitudinal Survey of Children and Youth (NLSCY)	2 (1.2%)
Project on Human Development in Chicago Neighborhoods (PHDCN)	2 (1.2%)
Twin and Offspring Study in Sweden (TOSS)	2 (1.2%)
Zurich Project on the Social Development of Children and Youths (Z-PROSO)	2 (1.2%)
Adachi Child Health Impact of Living Difficulty (A-CHILD)	1 (0.6%)
The Avon Longitudinal Study of Parents and Children (ALSPAC)	1 (0.6%)
Boricua Youth Study (BYS)	1 (0.6%)
Christchurch Health and Development Study (CHDS)	1 (0.6%)
Children of the National Longitudinal Survey of Youth (CNLSY)	1 (0.6%)
Cardiff Study of All Wales and North West of England Twins (CaStANET)	1 (0.6%)
Danish Mother of Twins Survey (DMTS)	1 (0.6%)
Danish National Birth Cohort (DNBC)	1 (0.6%)
Etude des Déterminants du développement et de la santé de l'eNfant (EDEN)	1 (0.6%)
Encuesta Longitudinal de la Primera Infancia cohort (ELPI)	1 (0.6%)
English and Romanian Adoptees (ERA)	1 (0.6%)
Fast Track Project (Fast Track Project)	1 (0.6%)
Finnish Medical Birth Register	1 (0.6%)
Healthy Babies Healthy Children (HBHC)	1 (0.6%)
Head Start Impact Study (HSIS)	1 (0.6%)
Iowa Adoptee Study	1 (0.6%)
Japanese study of Stratification, Health, Income, and Neighborhood (J-SHINE)	1 (0.6%)
Longitudinal Israeli Study of Twins (LIST)	1 (0.6%)
Longitudinal Study of Australian Children (LSAC)	1 (0.6%)
Midwest Infant Development Study (MIDS)	1 (0.6%)
Missouri Mothers and Their Children study (MO-MATCH)	1 (0.6%)
Missouri Adolescent Female Twin Study (MOAFTS)	1 (0.6%)

Characteristic	*k* (%)
National Comorbidity Survey Replication Adolescent Supplement (NCS-A)	1 (0.6%)
Nonshared Environment and Adolescent Development (NEAD)	1 (0.6%)
National Institute of Child Health and Human Development Study of Early Child Care and Youth Development (NICHD SECCYD)	1 (0.6%)
National Survey of Child and Adolescent Well-Being (NSCAW II)	1 (0.6%)
National Survey of Children's Health (NSCH)	1 (0.6%)
Ontario Child Health Study (OCHS)	1 (0.6%)
Pittsburgh Girls Study (PGS)	1 (0.6%)
Panel Study of Income Dynamics (PSID)	1 (0.6%)
Pittsburgh Youth Study (PYS)	1 (0.6%)
Quebec Longitudinal Study of Child Development (QLSCD)	1 (0.6%)
Sisters and Brothers Study (SBS)	1 (0.6%)
Swedish population-based registers (Swedish population-based registers)	1 (0.6%)
Twins and Multiple Births Association Heritability Study (TAMBAHS)	1 (0.6%)
Twin Study of Behavioral and Emotional Development in Children (TBED-C)	1 (0.6%)
Trondheim Early Secure Study (TESS)	1 (0.6%)
The Twins, Family and Behaviour (tFaB)	1 (0.6%)
Twin and Offspring Study in Sweden & the Twin Study of Child and Adolescent Development (TOSS & TCHAD)	1 (0.6%)
Toronto Sibling Study / Motherisk (Toronto Sibling Study / Motherisk)	1 (0.6%)
Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD)	1 (0.6%)
Wales Adoption Cohort Study (WACS)	1 (0.6%)
Women's Employment Study (WES)	1 (0.6%)
Young Parents and Their Children in Australia (YPCA)	1 (0.6%)
Young in Norway (Young in Norway)	1 (0.6%)
Not reported	11 (6.4%)

Abbreviations. k = number of studies, % = percentage.

3.3.3 Individual-level risk factors

I identified 11 different individual-level risk factors that have been studied using causal inference methods: prenatal exposure to toxins (*k* = 26), own education (*k* = 14), birth weight (*k* = 4), breastfeeding (*k* = 3), alcohol use (*k* = 3), maternal age at birth (*k* = 2), adolescent childbearing (*k* = 1), gaming (*k* = 1), obstetric complications (*k* = 1), tv viewing (*k* = 1) and temperament (*k* = 1).

3.3.3.1 Prenatal exposure to toxins

Twenty-six studies investigated the causal effect of prenatal exposure to substances and toxins on DBDs (Table **3.2**). These studies included data on 526,006 individuals from 18 distinct cohorts across eight countries. The studies examined seven types of substances and toxins, including: tobacco (*k* = 14), antidepressants (*k* = 4), alcohol (*k* = 2), paracetamol (*k* = 2), lead exposure (*k* = 1), nutrient deficiency (*k* = 1), general substances and toxins (*k* = 1) and triptans (*k* = 1). The most common causal inference methods used were the discordant sibling study design (*k* = 13), propensity score matching analyses (*k* = 6), the adoption study design (*k* = 3) and inverse probability of treatment weighting (*k* = 3; see Figure **3.3**).

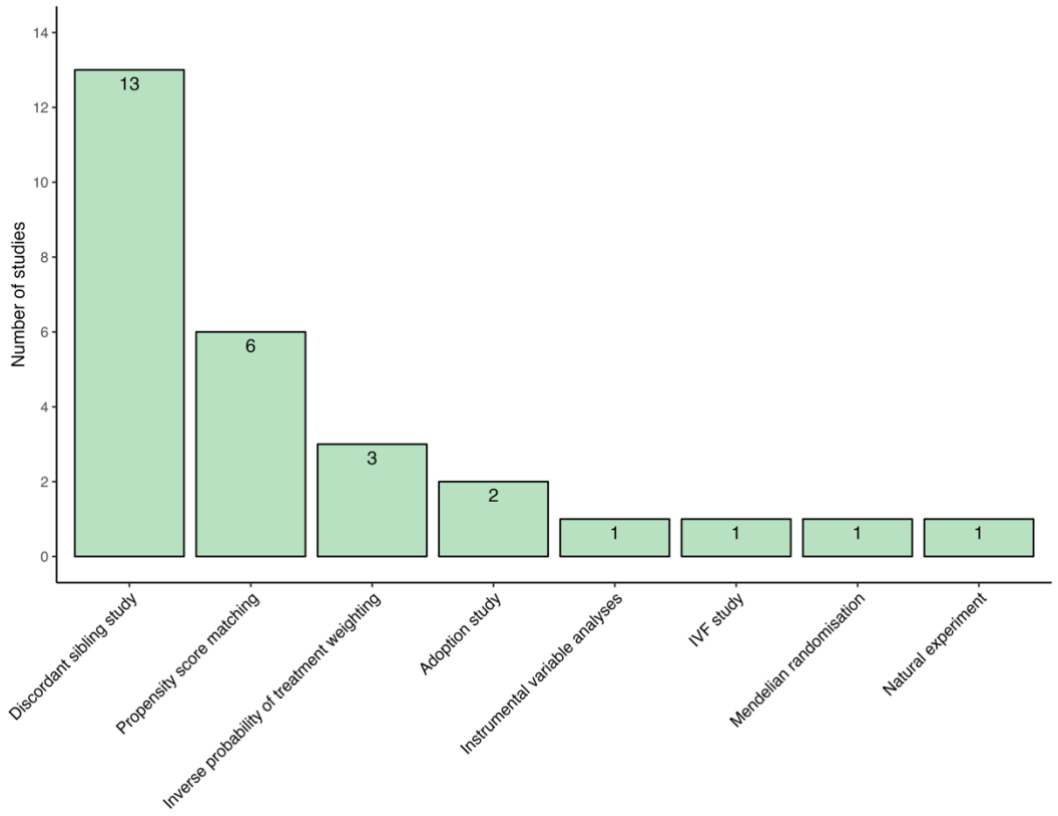


Figure 3.3 *The causal inference methods used in the studies that investigated the influence of prenatal exposure to substances and toxins on offspring disruptive behaviour disorder symptoms.*

Table 3.2 *Selected characteristics of studies using causal inference methods to investigate the influence of prenatal exposure to substances and toxins on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Boutwell *et al.,* (2010)	ECLS-B	USA	PSM	Smoking	EXT	3,343
Boutwell *et al.,* (2011)	FFCWS	USA	PSM	Smoking	EXT	1,951
Brandlistuen *et al.,* (2013)	MoBa	Norway	Sib	Paracetamol	EXT	1,878
Brandlistuen *et al.,* (2015)	MoBa	Norway	Sib	Anti-depressants	EXT	20,180
Brandlistuen *et al.,* (2017)	MoBa	Norway	Sib	Anti-depressants	EXT	38,594
D'Onofrio *et al.,* (2008)	NLSY79	USA	Sib	Smoking	CP	6,283
D'Onofrio *et al.,* (2012)	NLSY79	USA	Sib	Smoking	EXT	6,066
D'Onofrio, Van Hulle *et al.,* (2007)	NLSY79	USA	Sib	Alcohol	EXT	3,447

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Ekblad *et al.,* (2017)	Finnish Medical Birth Register	Finland	Sib	Smoking	EXT	300,336
Ekblad *et al.,* (2019)	MO-MATCH	USA	Sib	Smoking	EXT	346
Ellingson *et al.,* (2014)	NLSY79	USA	Sib	Smoking	EXT	10,251
Ellis *et al.,* (2012)	TESS	Norway	PSM	Smoking	ODD	995
Estabrook *et al.,* (2016)	MIDS	USA	Sib	Smoking	EXT	299
Gaysina, Fergusson *et al.,* (2013)	CHDS	New Zealand	Adopt	Smoking	CP	36
Gaysina, Fergusson *et al.,* (2013)	EGDS	USA	Adopt	Smoking	CP	311
Gaysina, Fergusson *et al.,* (2013)	C-IVF	UK	IVF	Smoking	CP	206
Gilman *et al.,* (2008)	CPP	USA	Sib	Smoking	CP	52,919
Harris *et al.,* (2018)	MoBa	Norway	IPTW	Triptans	EXT	37,656
McCrory *et al.,* (2012)	GUI	Ireland	PSM	Smoking	EXT	7,505
Murray *et al.,* (2016)	ALSPAC	UK	MR	Alcohol	Other	3,544
Neiderhiser *et al.,* (2016)	EGDS	USA	Adopt	Substances and toxins	EXT	561
Neugebauer *et al.,* (1999)	Not reported	The Netherlands	NE	Nutrient deficiency	ASPD	100,543
Nulman *et al.,* (2015)	Toronto Sibling Study / Motherisk	Canada	Sib	Anti-depressants	EXT	90
Palmer *et al.,* (2016)	MOAFTS	USA	PSM	Smoking	EXT	3,232
Paradis *et al.,* (2017)	CPP	USA	Sib	Smoking	Other	1,684
Sampson *et al.,* (2018)	PHDCN	USA	PSM	Lead exposure	EXT	1,255
Sampson *et al.,* (2018)	PHDCN	USA	IV	Lead exposure	EXT	1,255
Sundbakk *et al.,* (2019)	MoBa	Norway	IPTW	Anti-depressants	EXT	36,401
Tronnes *et al.,* (2019)	MoBa	Norway	IPTW	Paracetamol	EXT	32,934

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, DiD = difference in difference study, Sibling = discordant sibling study, Twin = discordant twin study, RG = fuzzy regression discontinuity design, IV = instrumental variable analyses, IPTW = inverse probability of treatment weighting, IVF = in-vitro fertilisation study, MR = Mendelian randomisation, NE = natural experiment, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, ODD = oppositional defiant disorder, Other = other disruptive behaviour disorder.

3.3.3.2 Own education

I identified 14 studies that examined the causal effect of education on DBDs using causal inference methods. They included data on 148,539 individuals from 14 distinct cohorts across nine countries. The most common causal inference methods were propensity score matching analyses (k = 5), instrumental variable analyses (k = 3), and discordant sibling study design (k = 2*;* Figure **3.4**). The studies investigated nine education measures, the majority of which focussed on childcare (Table **3.3**).

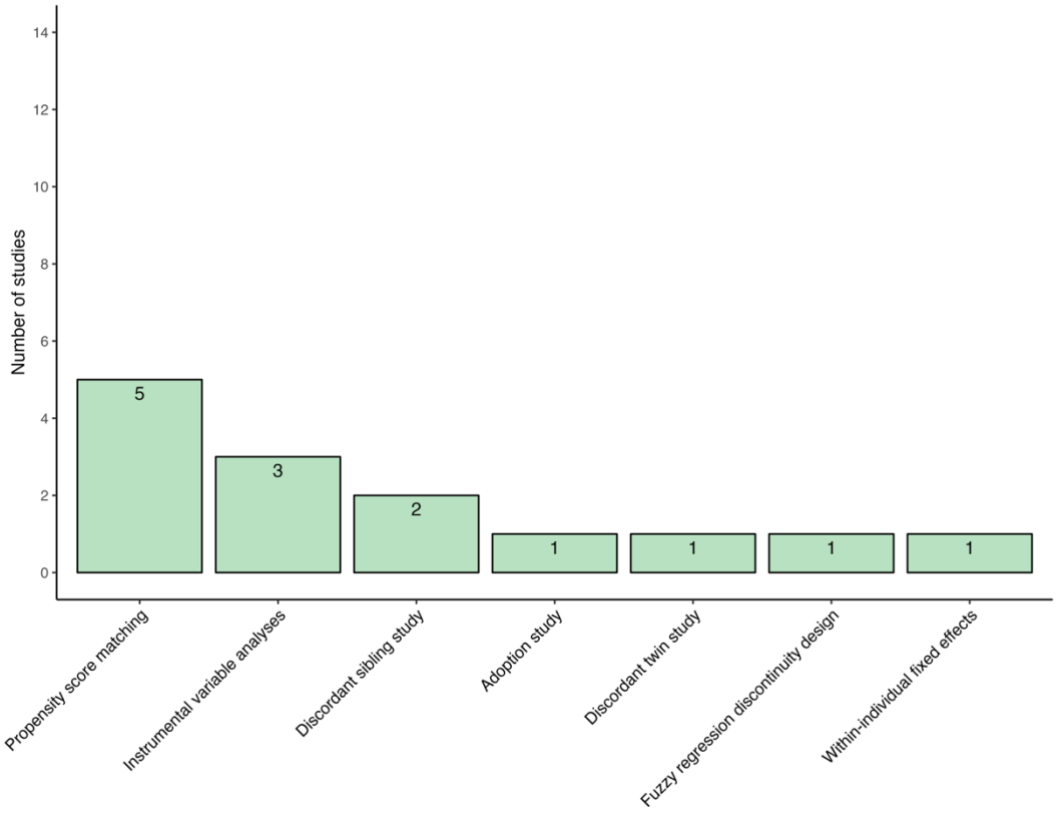


Figure 3.4 *The causal inference methods used in the studies that investigated the influence of education on offspring disruptive behaviour disorder symptoms.*

Table 3.3 *Selected characteristics of studies using causal inference methods to investigate the influence of education on disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Crosby *et al.,* (2010)	Six random-assignment studies	USA	IV	Childcare	EXT	3,290
Dee *et al.,* et al (2018)	DNBC	Denmark	RD	School starting age	CP	7,642
Edwards et al et al (2018)	YPCA	Australia	PSM	Childcare	EXT	317
Gomajee *et al.,* (2018)	EDEN	France	PSM	Childcare	Other	1,428
Herbst *et al.,* (2016)	ECLS-K	USA	IV	Childcare	EXT	3,848
Jaffee *et al.,* (2011)	CNLSY	USA	Sib	Childcare	CP	9,185
Lee *et al.,* (2018)	NICHD SECCYD	USA	FE	Unsupervised time with peers	EXT	747
Lee *et al.,* (2018)	NICHD SECCYD	USA	FE	Paid employment	EXT	747
Lee *et al.,* (2018)	NICHD SECCYD	USA	FE	Sports	EXT	747
Lee *et al.,* (2018)	NICHD SECCYD	USA	FE	Other organised activities after-school	EXT	747
Lipscomb *et al.,* (2014)	EGDS	USA	Adopt	Childcare	Other	233
Monnet *et al.,* (2019)	NSCH	USA	IV	Childcare	CP	42,462
Obsuth *et al.,* (2017)	Z-PROSO	Switzerland	PSM	Teacher-student relationship	Other	1,067
Oliver *et al.,* (2008)	TEDS	UK	Twin	Classroom environment	EXT	570
Orri *et al.,* (2019)	QLSCD	Canada	PSM	Childcare	CP	1,588
Powers *et al.,* (2016)	Fast Track Project	USA	PSM	Special education setting	CD	891
Zachrisson *et al.,* (2013)	MoBa	Norway	Sib	Childcare	EXT	75,271

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, DiD = difference in difference study, Sibling = discordant sibling study, Twin = discordant twin study, RG = fuzzy regression discontinuity design, IV = instrumental variable analyses, IPTW = inverse probability of treatment weighting, IVF = in-vitro fertilisation study, MR = Mendelian randomisation, NE = natural experiment, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, ODD = oppositional defiant disorder, Other = other disruptive behaviour disorder.

3.3.3.3 Birth weight

Four studies used the discordant twin study design to examine the causal effect of birth weight on DBDs. These studies included data on 10,122 individuals from 4 cohorts in 3 countries (Figure **3.5**; Table **3.4**)

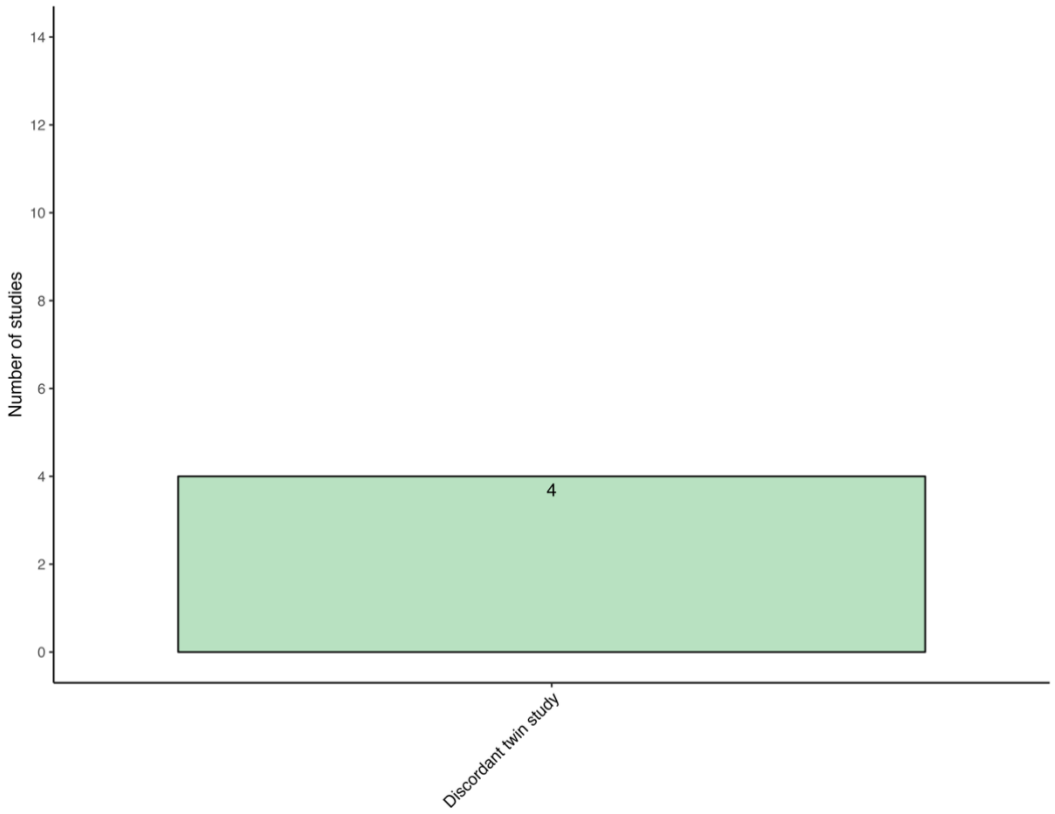


Figure 3.5 *The causal inference methods used in the studies that investigated the influence of birth weight on offspring disruptive behaviour disorder symptoms.*

Table 3.4 *Selected characteristics of studies using causal inference methods to investigate the influence of birthweight on disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Outcome	*n*
Asbury *et al.,* (2006)	TEDS	UK	Twin	Other	5,162
Mankuta *et al.,* (2010)	LIST	Israel	Twin	CP	224
Mollegaard *et al.,* (2020)	DMTS	Denmark	Twin	EXT	4,228
Tore *et al.,* (2018)	TAMBAHS	UK	Twin	EXT	508

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Twin = discordant twin study, Outcomes: CP = conduct problems, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.3.4 Breastfeeding

Three studies, which included data on 9,920 individuals (number of cohorts = 3; number of countries = 3), explored the potential causal effect of breastfeeding. Two studies used propensity score matching analyses, and the other used the IVF study design (Figure **3.6**; Table **3.5**).

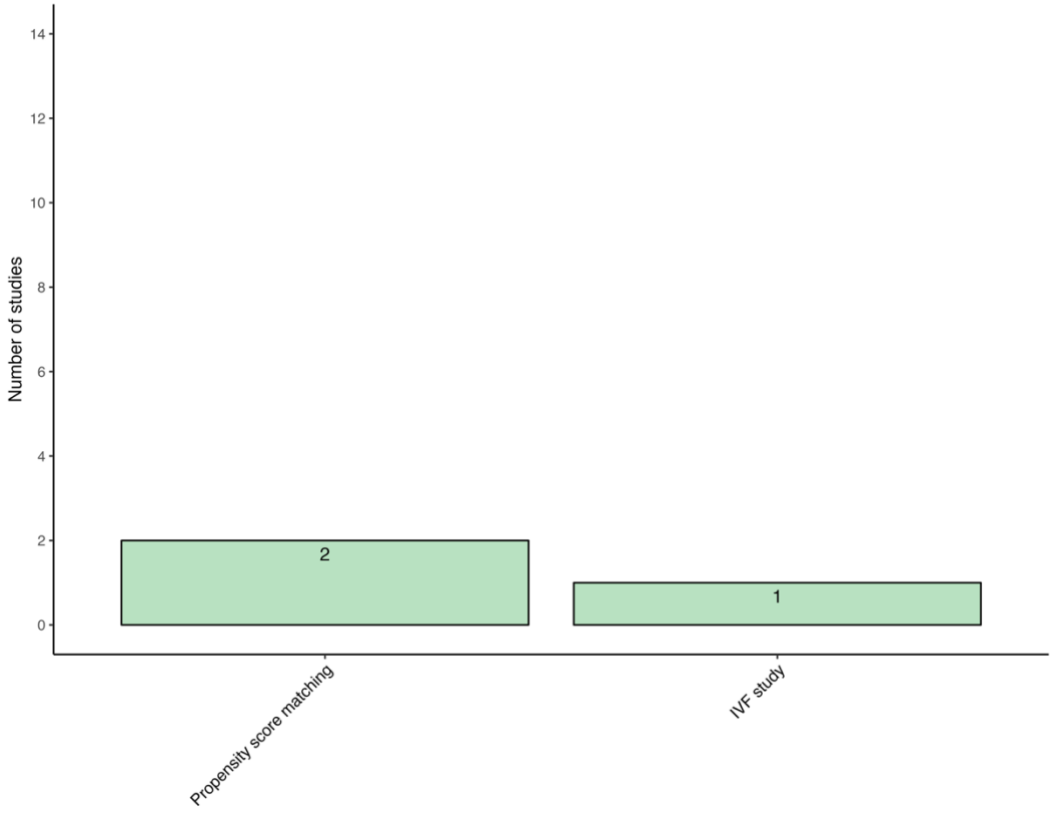


Figure 3.6 *The causal inference methods used in the studies that investigated the influence of breastfeeding on offspring disruptive behaviour disorder symptoms.*

Table 3.5 *Selected characteristics of studies using causal inference methods to investigate the influence of breastfeeding on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Outcome	*n*
Girard *et al.,* (2018)	GUI	Ireland	PSM	EXT	6,013
Girard *et al.,* (2019)	ELPI	Chile	PSM	EXT	3,037
Shelton *et al.,* (2011)	Not reported	USA/UK	IVF	CP	870

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: IVF = in-vitro fertilisation study, PSM = propensity score matching, Outcomes: CP = conduct problems, EXT = externalising symptoms.

3.3.3.5 Alcohol use

Three studies (*n* = 14,901 individuals, 3 cohorts, 3 countries) assessed the effect of alcohol use on disruptive behaviour using three causal inference methods: the discordant twin study design (*k* = 1), Mendelian randomisation analyses (*k* = 1) and within-individual fixed effects analyses (*k* = 1; Figure **3.7**; Table **3.6**)

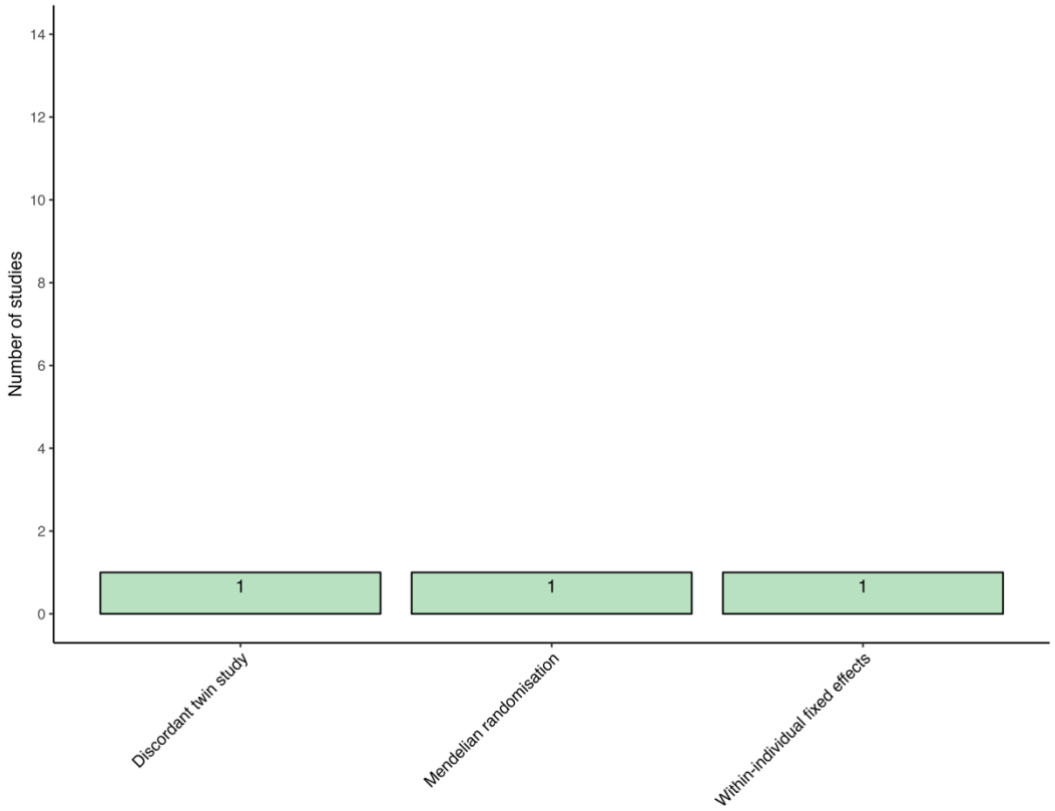


Figure 3.7 *The causal inference methods used in the studies that investigated the influence of substance use on offspring disruptive behaviour disorder symptoms.*

Table 3.6 *Selected characteristics of studies using causal inference methods to investigate the influence of substance use on disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Outcome	*n*
Chao *et al.,* (2017)	BeTwiSt	China	MR	EXT	1,608
Staff *et al.,* (2019)	MCS	UK	FE	EXT	10,529
Waldron *et al.,* (2018)	MTFS	USA	Twin	Other	2,764

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Twin = discordant twin study, MR = Mendelian randomisation, FE = within-individual fixed effects. Outcomes: EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.3.6 Other potential individual-level risk factors

Fewer than three studies examined the following risk factors: maternal age at birth (*k* = 2), adolescent childbearing (*k* = 1), gaming (*k* = 1), obstetric complications (*k* = 1), television viewing (*k* = 1) and temperament (Table **3.7**).

Table 3.7 *Selected characteristics of studies using causal inference methods to investigate other individual-level risk factors for disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Outcome	*n*
Maternal age at birth (*n* = 2)					
D'Onofrio *et al.,* (2009)	NLSY79	USA	Sib	CP	15,763
Harden, Lynch *et al.,* (2007)	OZALC	Australia	CoT	Other	1,364
Obstetric complications (*n* = 1)					
Neiderhiser *et al.,* (2016)	EGDS	USA	Adopt	EXT	561
Temperament (*n* = 1)					
Goodnight *et al.,* (2016)	NLSY79	USA	Sib	CP	9,237
Adolescent childbearing (*n* = 1)					
Hipwell *et al.,* (2016)	PGS	USA	PSM	CP	441
TV viewing (*n* = 1)					
Jackson *et al.,* (2018)	ECLS-B	USA	PSM	EXT	5,000
Gaming (*n* = 1)					
Brunborg *et al.,* (2014)	Young in Norway	Norway	FE	CP	1,928

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, Sib = discordant sibling study, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: CP = conduct problems, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4 Familial-level risk factors

I identified nine familial-level risk factors that have been examined using causal inference methods: parenting practices (*k* = 39), parental internalising symptoms (*k* = 17), adverse childhood experiences (*k* = 14), family dynamics (*k* = 13), parental substance use (*k* = 10), parental externalising symptoms (*k* = 7), family socioeconomic position (*k* = 6), parental education and employment (*k* = 5) and residential mobility (*k* = 1).

3.3.4.1 Parenting practices

Parenting practices were the most studied risk factor using causal inference methods. Thirty-nine studies used eight different causal inference methods to investigate the causal effect of parenting practices on DBDs (Table **3.8**). These studies included data on 57,535 individuals from 24 distinct cohorts in 7 countries. The most common causal inference methods used were the adoption study design (*k* = 14), the discordant twin study design (*k* = 11) and the discordant sibling study design (*k* = 9; Figure **3.8**).

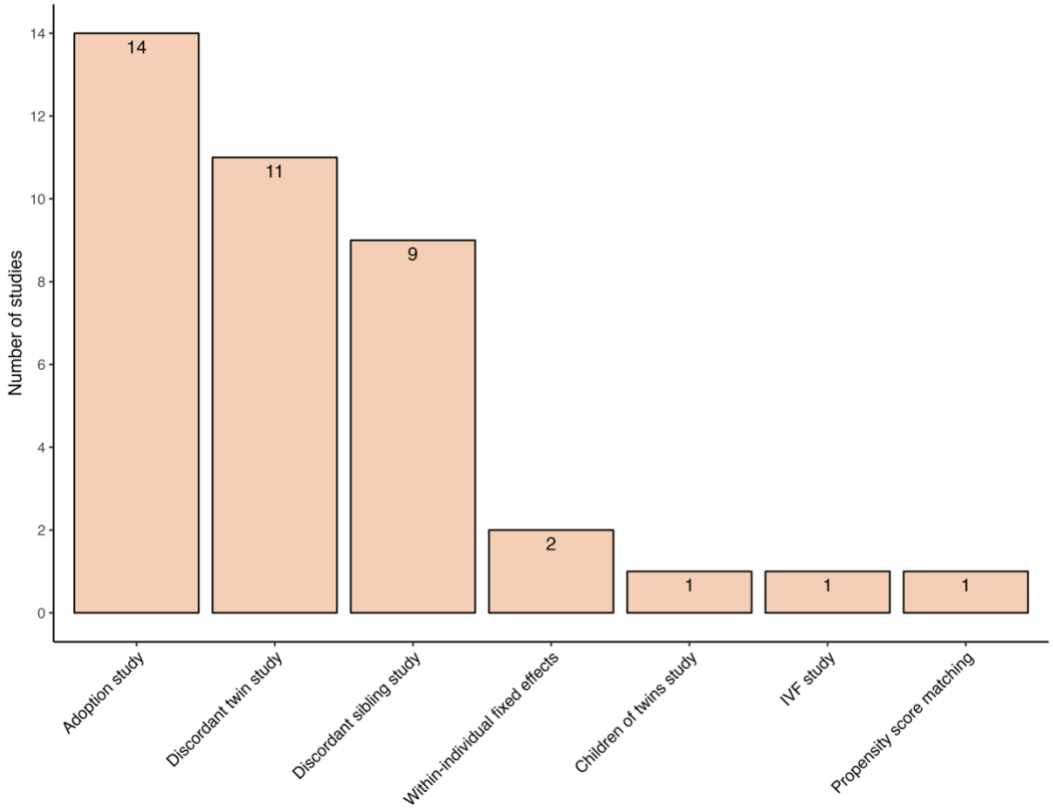


Figure 3.8 *The causal inference methods used in the studies that investigated the influence of parenting practices on offspring disruptive behaviour disorder symptoms.*

Table 3.8 *Selected characteristics of studies using causal inference methods to investigate the influence of parenting practices on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Anthony *et al.,* (2019)	WACS	UK	Adopt	Parental Warmth	EXT	62
Asbury *et al.,* (2003)	TEDS	UK	Twin	Harsh Discipline; Parental Feeling	Other	4,706
Asbury *et al.,* (2006)	TEDS	UK	Twin	Parent-Child Communication; Parental Feeling; Harsh Discipline	Other	5,162
Barnett *et al.,* (2013)	FFCWS	USA	Sib	Parental Warmth; Coercive Parenting	EXT	274

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Besemer *et al.,* (2016)	PYS	USA	FE	Harsh Discipline; Parental Involvement; Parent-Child Communication	CP or ODD	487
Boisvert *et al.,* (2008)	PSID	USA	Sib	Parental Warmth; Parental Monitoring	EXT	1,759
Boyle *et al.,* (2004)	OCHS; NLSCY; NLSY79	Canada	Sib	Parental Involvement; Parental Warmth; Harsh Discipline; Parental Hostility; Physical Discipline	EXT	2,128
Caspi *et al.,* (2004)	E-Risk	UK	Twin	Expressed Emotion; Parental Warmth; Parental Criticism	EXT	7,392
Cecil *et al.,* (2012)	TEDS	UK	Twin	Harsh Discipline; Parental Feeling	EXT	2,876
Deater-Deckard *et al.,* (2004)	N2CAP	USA	Adopt	Parent-Child Relationship	CD or ODD	1,506
Glover *et al.,* (2010)	N2CAP	USA	Adopt	Expressed Emotion; Parental Feeling	EXT	1,244
Harold *et al.,* (2012)	CardiffIVF	UK/USA	IVF	Parental Hostility	CP	5,184
Harold *et al.,* (2013)	EGDS	USA	Adopt	Parental Hostility	EXT	396
Hou *et al.,* (2013)	BeTwiSt	China	Twin	Parental Warmth; Parental Hostility	EXT	85
Klahr, McGue *et al.,* (2011)	SIBS	USA	Adopt	Parent-Child Conflict	CP	377
Klahr, Rueter *et al.,* (2011)	SIBS	USA	Adopt	Parent-Child Conflict; Coercive Parenting	EXT	218
Latham *et al.,* (2017)	TFaB	UK	Sib	Coercive Parenting	EXT	1,040

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Leve *et al.,* (2009)	EGDS	USA	Adopt	Structured	CP or ASPD	672
Lipscomb *et al.,* (2014)	EGDS	USA	Adopt	Overreactive Parenting	Other	390
Long *et al.,* (2015)	VATSPSUD	USA	Twin	Low Parental Warmth; Overprotective Parenting; Harsh Discipline	EXT	212
Lysenko *et al.,* (2013)	TEDS	UK	Sib	Harsh Discipline	EXT	290
Marceau *et al.,* (2013)	EGDS	USA	Adopt	Overreactive Parenting	Other	233
Mark *et al.,* (2017)	SBS	UK	Sib	Parent-Child Relationship; Parent-Child Conflict	CD	2,606
Meunier, Bisceglia *et al.,* (2012)	HBHC	Canada	Sib	Parent-Child Relationship	CP	27,660
Morcillo *et al.,* (2011)	BYS	USA/Puerto Rico	PSM	Family Bonding	EXT	561
Narusyte *et al.,* (2011)	TOSS	Sweden	CoT	Parental Criticism	CP	156
Oliver *et al.,* (2015)	TEDS	UK	Twin	Parental Feeling	EXT	809
Pike *et al.,* (1996)	NEAD	USA	Twin	Parent-Child Conflict	Other	2,491
Reuben *et al.,* (2016)	EGDS	USA	Adopt	Parental Warmth, Overreactive Parenting	EXT	1,818
Richmond *et al.,* (2006)	Not reported	Not reported	Sib	Parental Hostility	CP	6,308
Richmond *et al.,* (2009)	Not reported	Not reported	Sib	Parental Hostility	Other	93
Riggins-Caspers *et al.,* (2003)	Not reported	Not reported	Adopt	Physical Discipline; Harsh Discipline	EXT	225
Rolon-Arroyo *et al.,* (2018)	Not reported	USA	FE	Overreactive Parenting	EXT	186
Roos *et al.,* (2016)	EGDS	USA	Adopt	Parental Involvement	Other	228
Samek *et al.,* (2014)	SIBS	USA	Adopt	Parental Involvement; Parent-Child Conflict; Parent-Child Relationship	CD	150

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Shelton *et al.,* (2008)	CaStANET	UK	Twin	Parental Hostility; Parental Warmth	CD or ODD	199
Viding *et al.,* (2009)	TEDS	UK	Twin	Harsh Discipline	EXT	293
Waller, Hyde *et al.,* (2018)	TBED-C	USA	Twin	Parental Warmth; Harsh Discipline	ASPD	533
Anthony *et al.,* (2019)	WACS	UK	Adopt	Parental Warmth	CP	462
Asbury *et al.,* (2003)	TEDS	UK	Twin	Harsh Discipline; Parental Feeling	CP	4,508
Asbury *et al.,* (2006)	TEDS	UK	Twin	Parent-Child Communication; Parental Feeling; Harsh Discipline	CP	454

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, DiD = difference in difference study, Sibling = discordant sibling study, Twin = discordant twin study, RG = fuzzy regression discontinuity design, IV = instrumental variable analyses, IPTW = inverse probability of treatment weighting, IVF = in-vitro fertilisation study, MR = Mendelian randomisation, NE = natural experiment, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, ODD = oppositional defiant disorder, Other = other disruptive behaviour disorder.

3.3.4.2 Parental internalising

I identified 17 studies that used seven causal inference methods to investigate the impact of parental internalising symptoms on disruptive behaviour (Table **3.9**). These studies included data on 190,556 individuals from eight cohorts across five countries. The six types of parental internalising symptoms included: depression (*k* = 12), stress (*k* = 2), anxiety (*k* = 1), depression and anxiety (*k* = 1), broad internalising (*k* = 1) and post-natal depression (*k* = 1). The most common causal inference methods used were the adoption study design (*k* = 9), the children of twins study design (*k* = 3) and the discordant sibling study design (*k* = 3; Figure **3.9**).

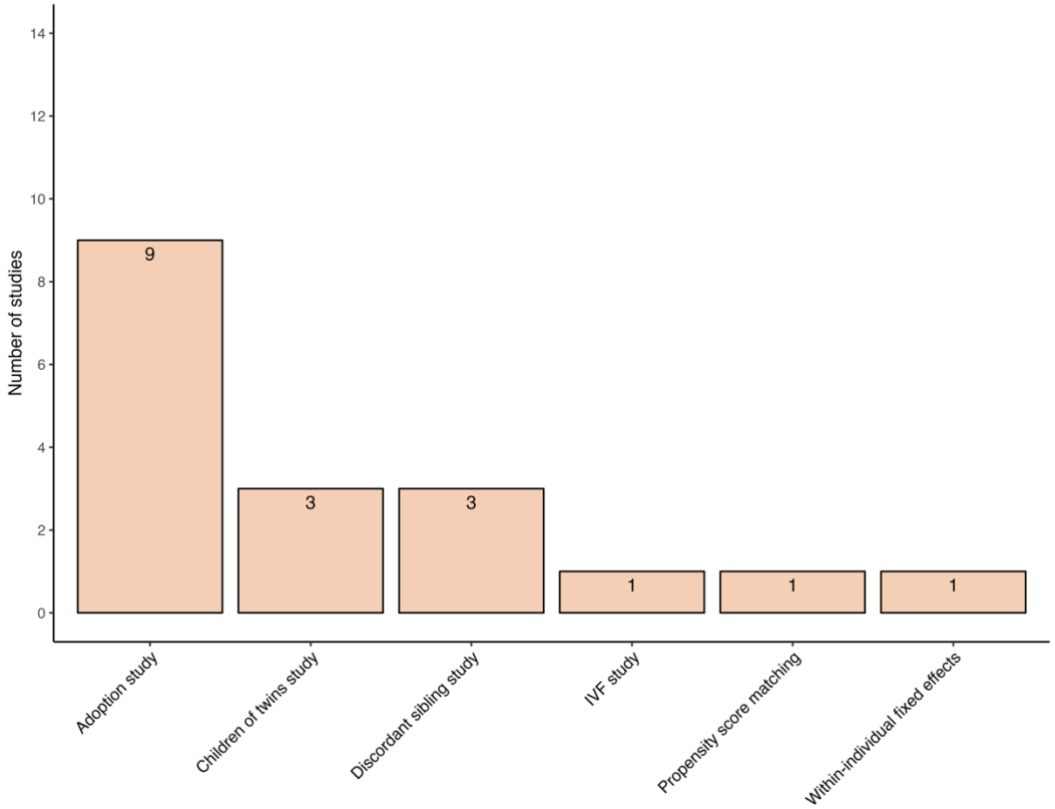


Figure 3.9 *The causal inference methods used in the studies that investigated the influence of parental internalising symptoms on offspring disruptive behaviour disorder symptoms.*

Table 3.9 *Selected characteristics of studies using causal inference methods to investigate the influence of parental internalising symptoms on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Gjerde *et al.,* (2017)	MoBa	Norway	Sib	Depression	EXT	17,830
Gjerde *et al.,* (2020)	MoBa	Norway	Sib	Anxiety	EXT	17,724
Grabow *et al.,* (2017)	EGDS	USA	Adopt	Depression	EXT	541
Hails *et al.,* (2019)	EGDS	USA	Adopt	Depression	EXT	503
Hannigan *et al.,* (2018)	MoBa	Norway	CoT	Post-natal depression	EXT	35,299
Kendler *et al.,* (2019)	Swedish population-based registers	Sweden	Sib	Depression	CD	146,216
Kerr *et al.,* (2013)	EGDS	USA	Adopt	Depression	Other	346
King *et al.,* (2018)	FFCWS	USA	FE	Depression	EXT	2,044
King *et al.,* (2018)	FFCWS	USA	FE	Stress	EXT	2,044
Marceau *et al.,* (2013)	EGDS	USA	Adopt	Depression and anxiety	EXT	561
McAdams *et al.,* (2015)	TOSS	Sweden	CoT	Depression	EXT	1,752
McAdams *et al.,* (2015)	EGDS	USA	Adopt	Depression	EXT	361
Pemberton *et al.,* (2010)	EGDS	USA	Adopt	Depression	EXT	351
Rice *et al.,* (2010)	C-IVF	UK	IVF	Stress	Other	474
Roos *et al.,* (2016)	EGDS	USA	Adopt	Internalising	EXT	293
Singh *et al.,* (2011)	ATR	Australia	CoT	Depression	CD	2,554
Taraban *et al.,* (2019)	EGDS	USA	Adopt	Depression	EXT	1,038
Tully *et al.,* (2008)	SIBS	USA	Adopt	Depression	CD	568
Turney *et al.,* (2012)	FFCWS	USA	PSM	Depression	DBD	2,655

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, Sib = discordant sibling study, IVF = in-vitro fertilisation study, NE = natural experiment, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: CD = conduct disorder, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4.3 Family dynamics

Thirteen studies examined the impact of negative family dynamics on DBDs (Table **3.10**). These studies used seven causal inference methods and included data on 24,547 individuals (n = 8 cohorts; 5 countries). The studies investigated four types of negative family dynamics, including parental separation (k = 8), marital conflict (k = 3), family functioning (k = 1) and parental relationship quality (k = 1). The most common methods were the children of twins study design (k = 3), the discordant sibling study design (k = 3) and within-individual fixed effects analyses (k = 3; Figure **3.10**).

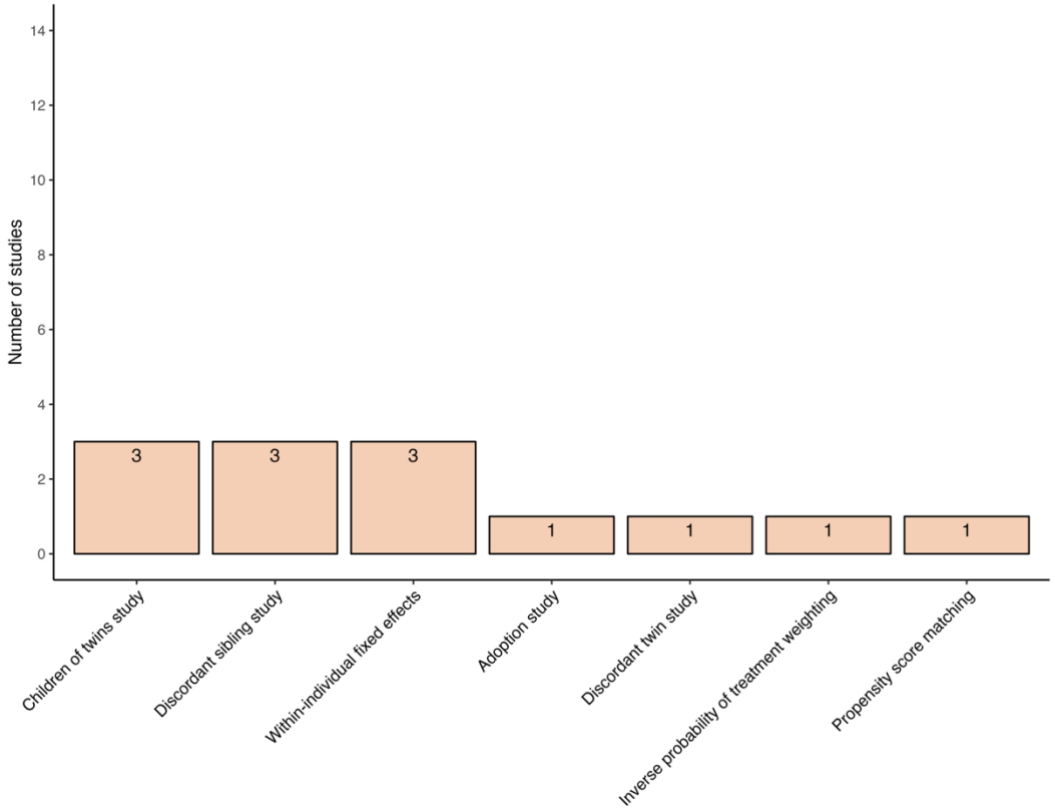


Figure 3.10 *The causal inference methods used in the studies that investigated the influence of family dynamics on offspring disruptive behaviour disorder symptoms.*

Table 3.10 *Selected characteristics of studies using causal inference methods to investigate the influence of negative family dynamics on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Burt *et al.,* (2008)	SIBS	USA	Adopt	Separation	Other	204
Burt *et al.,* (2010)	MTFS	USA	Twin	Separation	ASPD	578
D'Onofrio *et al.,* (2005)	OZALC	Australia	CoT	Separation	EXT	2,554
Fitzsimons *et al.,* (2019)	MCS	UK	FE	Separation	CP	6,245
Goldberg *et al.,* (2014)	FFCWS	USA	FE	Parental relationship quality	EXT	773
Harden, Turkheimer *et al.,* (2007)	OZALC	Australia	CoT	Marital conflict	CP	1,131
Lee *et al.,* (2015)	FFCWS	USA	IPTW	Separation	EXT	2,952
Mostafa *et al.,* (2018)	MCS	UK	FE	Separation	EXT	14,833
Richmond *et al.,* (2003)	NR	USA	Sib	Marital conflict	EXT	122
Richmond *et al.,* (2009)	NR	Not reported	Sib	Separation	Other	228
Schermerhorn *et al.,* (2011)	TOSS	Sweden	CoT	Family functioning	EXT	1,818
Skopp *et al.,* (2005)	NR	USA	Sib	Marital conflict	EXT	244
Weaver *et al.,* (2015)	NICHD	USA	PSM	Separation	EXT	1,364

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, Sib = discordant sibling study, Twin = discordant twin study, IPTW = inverse probability of treatment weighting, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CP = conduct problems, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4.4 Parental substance use

Ten studies (*k* = 204,002 individuals; 5 cohorts; 4 countries; Table **3.11**) investigated the influence of parental substance (ab)use, including alcohol (*k* = 3), drugs (*k* = 2), tobacco (*k* = 2) and cannabis (*k* = 1). The studies used six causal inference methods, the most common of which were the children of twins (*k* = 3) and adoption study designs (*k* = 3; Figure **3.11**)

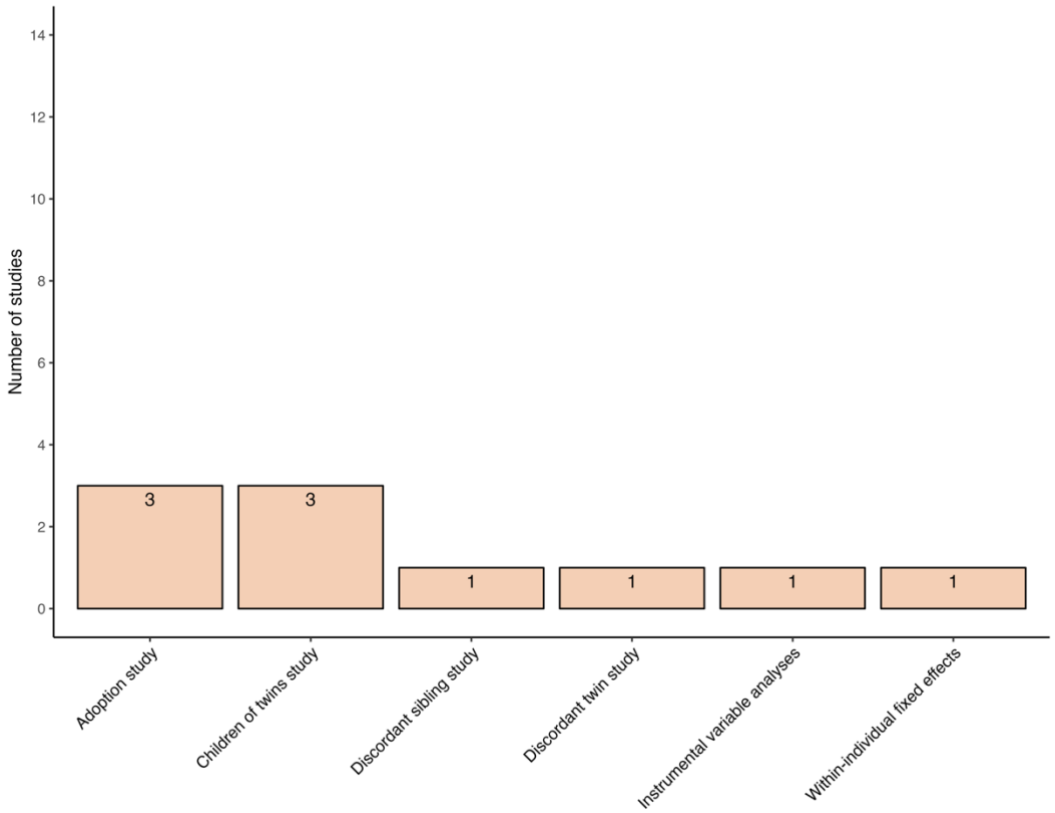


Figure 3.11 *The causal inference methods used in the studies that investigated the influence of parental substance use on offspring disruptive behaviour disorder symptoms.*

Table 3.11 *Selected characteristics of studies using causal inference methods to investigate the influence of parental substance use on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Haber *et al.,* (2005)	VET	USA	CoT	Alcohol	CD	1,270
Haber *et al.,* (2010)	VET	USA	CoT	Alcohol and drugs	CD	1,917
Kendler *et al.,* (2019)	Swedish population-based registers	Sweden	Sib	Alcohol and drugs	CD	146,216
Keyes, Legrand *et al.,* (2008)	SIBS	USA	Adopt	Smoking	Other	785
King *et al.,* (2009)	SIBS	USA	Adopt	Alcohol	Other	525
Knudsen *et al.,* (2015)	MoBa	Norway	FE	Alcohol	EXT	51,115
Lund *et al.,* (2019)	MoBa	Norway	IV	Alcohol	EXT	25,744
Samek *et al.,* (2014)	SIBS	USA	Adopt	Alcohol and smoking	ASPD	533
Scherrer *et al.,* (1996)	VET	USA	Twin	Cannabis	CD or ASPD	3,394
Waldron *et al.,* (2009)	OZALC	Australia	CoT	Alcohol	EXT	2,492

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, Sib = discordant sibling study, Twin = discordant twin study, IV = instrumental variable analyses, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CD = conduct disorder, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4.5 Adverse childhood experiences

I identified 12 studies that used five causal inference methods to investigate the effect of adverse childhood experiences (ACEs) on DBDs (Figure **3.12**). These studies included data on 32,224 individuals from 8 cohorts across five countries (Table **3.12**). Seven types of ACEs were included: physical abuse (k = 3), sexual abuse (k = 3), various ACEs (k = 1), deprivation (k = 1), witnessing intimate partner violence (k = 1), neglect (k = 1) and combined sexual abuse and physical maltreatment (k = 1). These were examined using propensity score matching analyses (k = 5), the discordant twin study design (k = 3) and within-individual fixed effects (k = 3).

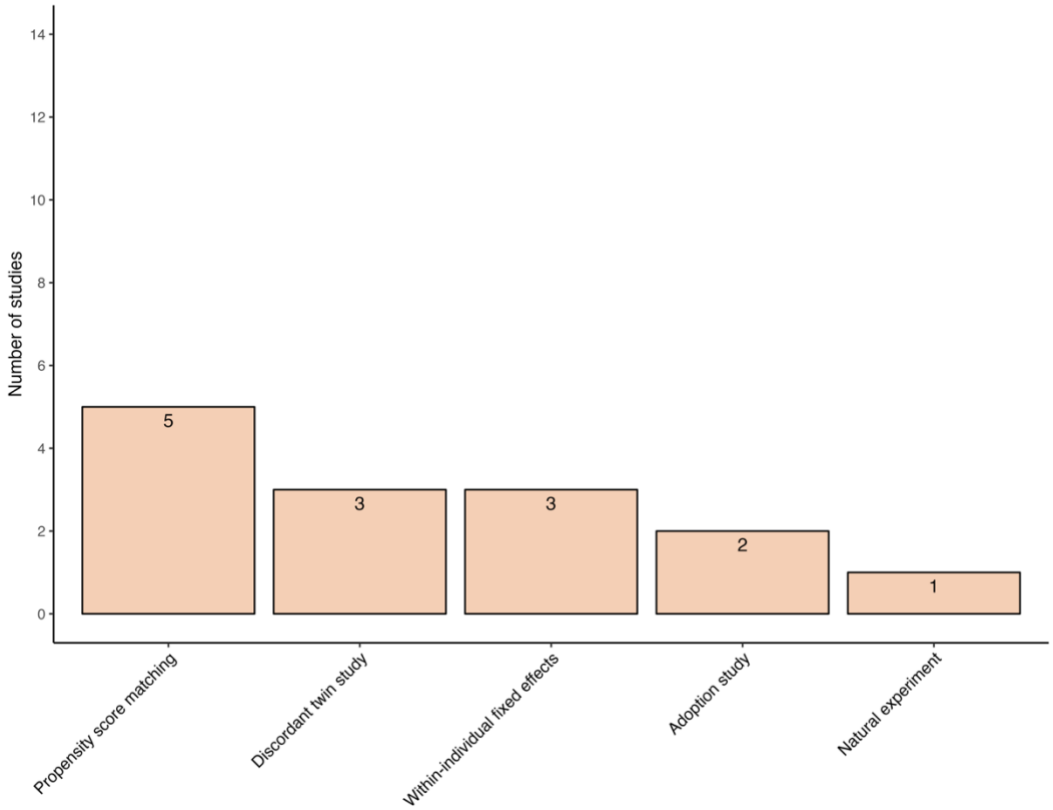


Figure 3.12 *The causal inference methods used in the studies that investigated the influence of adverse childhood experiences on disruptive behaviour disorder symptoms.*

Table 3.12 *Selected characteristics of studies using causal inference methods to investigate the influence of adverse childhood experiences on disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Averdijk *et al.,* (2018)	Z-PROSO	Switzerland	PSM	Foster care	EXT	1483
Beach *et al.,* (2013)	Iowa Adoptee Study	USA	Adopt	Sexual abuse	ASPD	155
Cadoret *et al.,* (1995)	NR	Not reported	Adopt	ACEs	CD	197
Dinwiddie *et al.,* (2000)	ATR	Australia	Twin	Sexual abuse	CP	2,682
Doi *et al.,* (2018)	A-CHILD	Japan	PSM	Neglect	EXT	4,195
Emery *et al.,* (2011)	PHDCN	USA	FE	Intimate partner violence	EXT	1,816
Gershoff *et al.,* (2018)	ECLS-K	USA	PSM	Physical abuse	EXT	12,112
Ma *et al.,* (2018)	FFCWS	USA	FE	Physical abuse	Other	2,472
Ma *et al.,* (2020)	FFCWS	USA	FE	Physical abuse	EXT	2,472
Misheva *et al.,* (2017)	ATR	Australia	Twin	Sexual abuse and physical maltreatment	CD	11,060
Nelson *et al.,* (2002)	ATR	Australia	Twin	Sexual abuse	CD	1,991
Sonuga-Barke *et al.,* (2017)	ERA	UK	NE	Deprivation	Other	217
Williams *et al.,* (2016)	ECLS-K	USA	PSM	Parental bereavement	EXT	250
Wu *et al.,* (2015)	NSCAW II	USA	PSM	Foster care	EXT	1,054

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, Twin = discordant twin study, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4.6 Parental externalising symptoms

Seven studies examined the causal effect of parental externalising symptoms on offspring DBDs, including parental incarceration (*k* = 4), parental antisocial behaviour symptoms (*k* = 2) and parental conduct disorder symptoms (*k* = 1). These studies included 6,480 individuals from five cohorts in four countries (Table **3.13**) and used four causal inference methods: propensity score matching analyses (*k* = 4), the adoption study design (*k* = 1), the children of twins study design (*k* = 1) and the IVF study design (*k* = 1; Figure **3.13**)

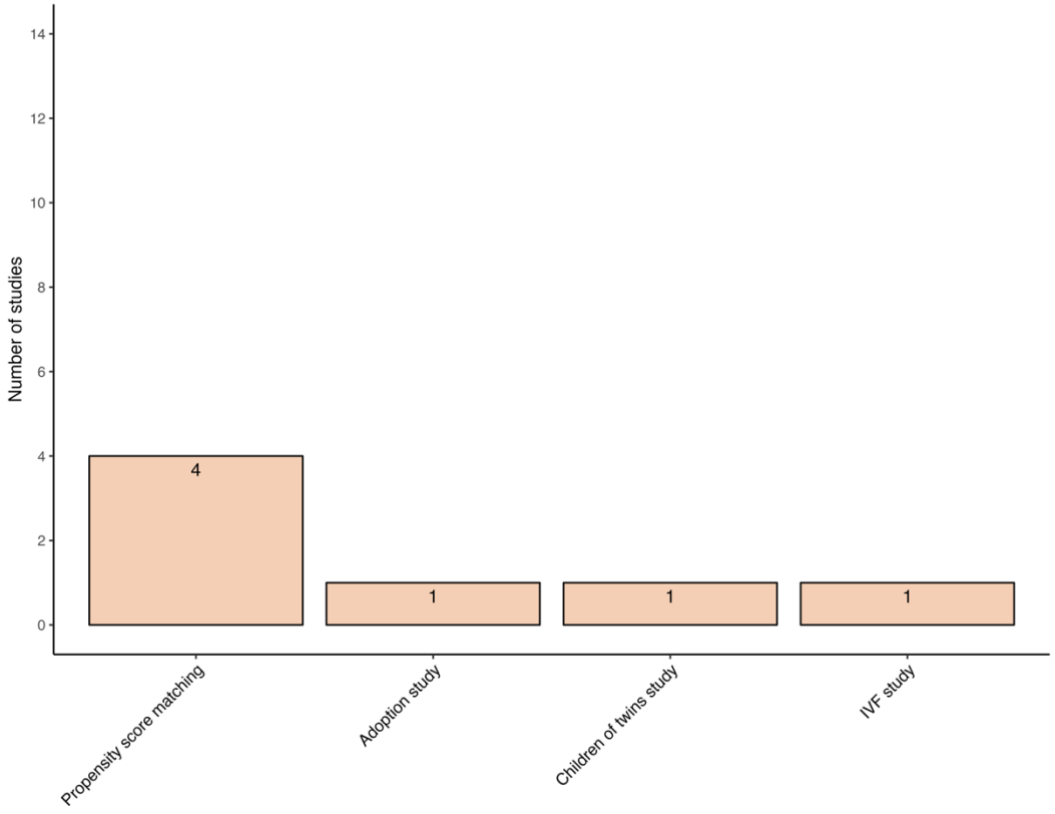


Figure 3.13 *The causal inference methods used in the studies that investigated the influence of parental externalising symptoms on offspring disruptive behaviour disorder symptoms.*

Table 3.13 *Selected characteristics of studies using causal inference methods to investigate the influence of parental externalising symptoms on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Bradshaw *et al.,* (2020)	GUI	Ireland	PSM	Incarceration	EXT	100
Copp *et al.,* (2018)	FFCWS	USA	PSM	Incarceration	EXT	3,196
D'Onofrio, Slutske *et al.,* (2007)	ATR	Australia	CoT	Conduct disorder	CD	2,554
Harold *et al.,* (2011)	C-IVF	UK	IVF	Antisocial behaviour	Other	283
Haskins *et al.,* (2015)	FFCWS	USA	PSM	Incarceration	EXT	2,162
Kerr *et al.,* (2013)	EGDS	USA	Adopt	Antisocial behaviour	Other	347
Turney *et al.,* (2017)	FFCWS	USA	PSM	Incarceration	EXT	3,065

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, IVF = in-vitro fertilisation study, PSM = propensity score matching. Outcomes: CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4.7 Socioeconomic position

I identified six studies that used four different causal inference methods to investigate the impact of socioeconomic position on disruptive behaviour (*n* = 13,972 individuals; 4 cohorts; 2 countries; Table **3.14**). These studies used the discordant sibling study design (*k* = 2), the natural experiment study design (*k* = 2), the discordant twin study design (*k* = 1) and within-individual fixed effects analyses (*k* = 1; Figure **3.14**).

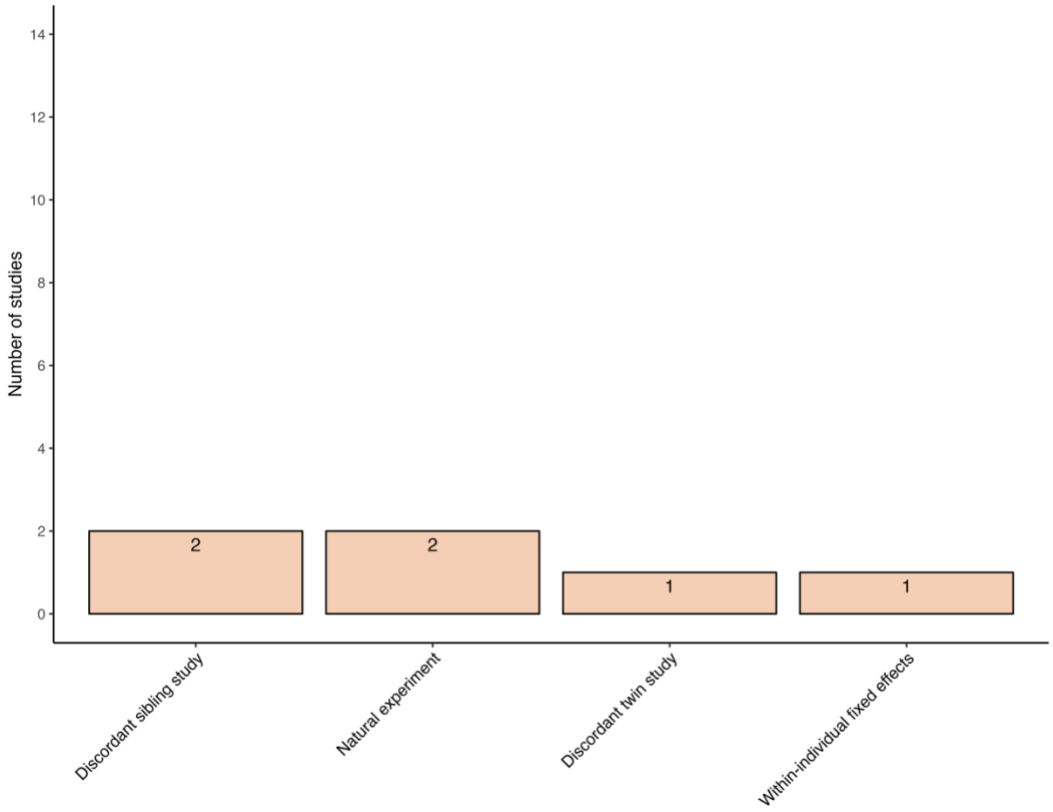


Figure 3.14 *The causal inference methods used in the studies that investigated the influence of familial Socioeconomic position on offspring disruptive behaviour disorder symptoms.*

Table 3.14 *Selected characteristics of studies using causal inference methods to investigate the influence of familial Socioeconomic position on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Costello *et al.,* (2003)	GSMS	USA	NE	Socioeconomic position	CD or ODD	1,420
Costello *et al.,* (2010)	GSMS	USA	NE	Socioeconomic position	CD, ODD or ASPD	1,420
D'Onofrio *et al.,* (2009)	NLSY79	USA	Sib	Socioeconomic position	CP	4,912
King *et al.,* (2018)	FFCWS	USA	FE	Poverty	EXT	2,044
Ramanathan *et al.,* (2017)	NLSY79	USA	Sib	Socioeconomic position	EXT	8,276
Rivenbark *et al.,* (2020)	E-Risk	UK	Twin	Socioeconomic position	CP	2,232

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Sib = discordant sibling study, Twin = discordant twin study, NE = natural experiment, FE = within-individual fixed effects. Outcomes: ASPD = antisocial personality disorder, CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, ODD = oppositional defiant disorder.

3.3.4.8 Parental education and employment

Five studies focused on the influence of parental education and employment on offspring disruptive behaviour and included 13,107 individuals from five distinct cohorts across three countries (Table **3.15**). The types of measures included parental education (*k* = 1), intensity of parental employment (*k* = 1), length of maternity leave (*k* = 1), parental employment status (*k* = 1) and type of parental employment (*k* = 1). The causal inference methods used included propensity score matching analyses (*k* = 3), the adoption study design (*k* = 1) and within-individual fixed effects analyses (*k* = 1; Figure **3.15**).

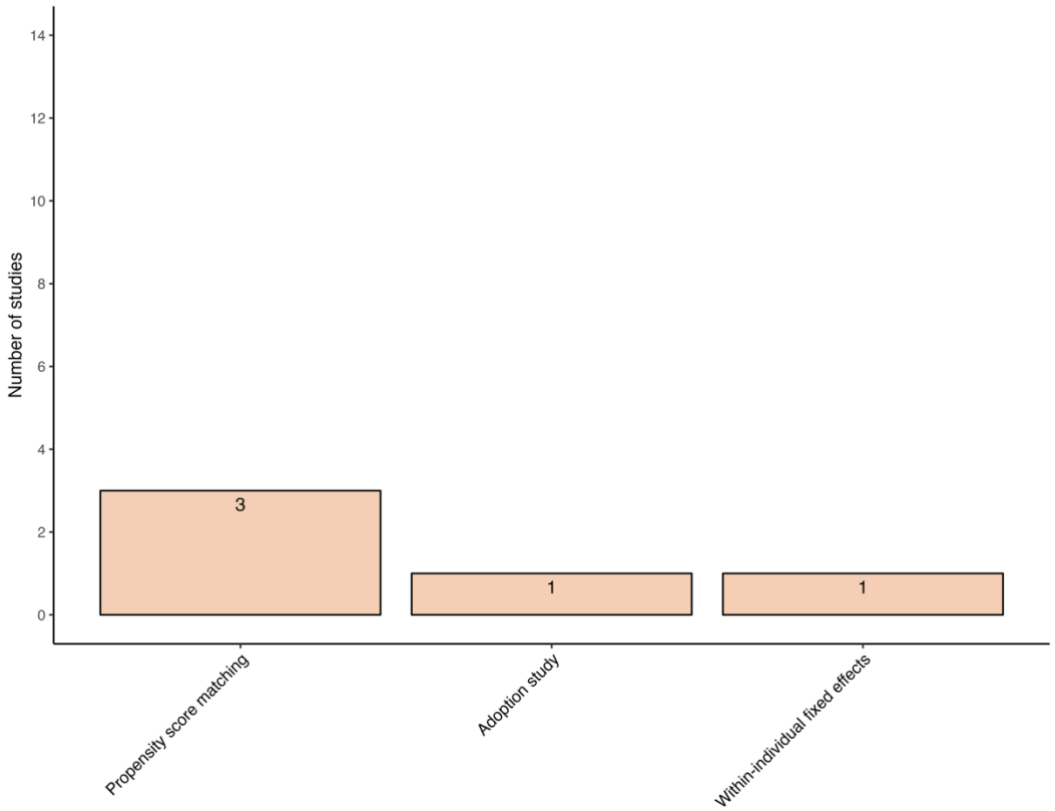


Figure 3.15 *The causal inference methods used in the studies that investigated the influence of parental employment on offspring disruptive behaviour disorder symptoms.*

Table 3.15 *Selected characteristics of studies using causal inference methods to investigate the influence of parental employment on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Berger *et al.,* (2005)	NLSCY	USA	PSM	Maternity leave	EXT	769
Dunifon *et al.,* (2003)	WES	USA	FE	Status	EXT	573
Duyme *et al.,* (1990)	NR	France	Adopt	Type	Other	77
Lombardi *et al.,* (2014)	HSIS	UK	PSM	Education	EXT	1,588
Harding *et al.,* (2015)	ECLS–B	USA	PSM	Intensity	CP	10,100

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.4.9 Parental adverse childhood experiences

Three studies examined the influence of parental exposure to ACEs on offspring DBDs using data on 745 individuals from 2 cohorts in the United States (Table **3.16**). Two of these studies used propensity score matching analyses, and the other used the adoption study design (Figure **3.16**).

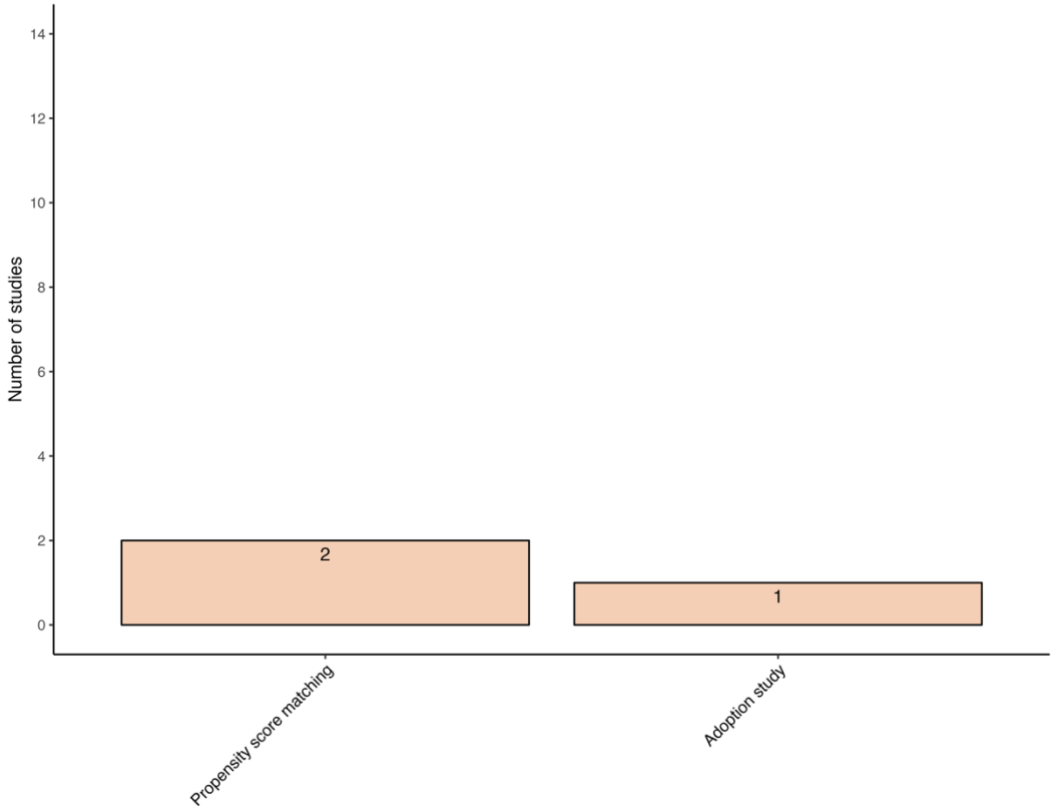


Figure 3.16 *The causal inference methods used in the studies that investigated the influence of parental adverse childhood experiences on offspring disruptive behaviour disorder symptoms.*

Table 3.16 *Selected characteristics of studies using causal inference methods to investigate the influence of parental adverse childhood experiences on offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Grabow *et al.,* (2017)	EGDS	USA	Adopt	Trauma	EXT	541
Zvara *et al.,* (2017)	FLP	USA	PSM	Sexual abuse	CP	204
Zvara *et al.,* (2019)	FLP	USA	PSM	Sexual abuse	CP	204

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, PSM = propensity score matching. Outcomes: CP = conduct problems, EXT = externalising symptoms.

3.3.4.10 Other potential familial risk factors

One study (*n* = 1,056 individuals) used propensity score matching analyses to examine the impact of residential mobility on disruptive behaviour (Table **3.17**).

Table 3.17 *Selected characteristics of studies using causal inference methods to investigate other familial-level risk factors for disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Outcome	*n*
Anderson *et al.,* (2017)	NICHD	USA	PSM	EXT	1,056

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: PSM = propensity score matching. Outcomes: EXT = externalising symptoms.

3.3.5 Extra-familial level risk factors

3.3.5.1 Neighbourhood characteristics

Seven studies using three different causal inference methods (within-individual fixed effects analyses, *k* = 4; propensity score matching analyses, *k* = 2; difference-in-difference, *k* = 1) examined the effect of neighbourhood characteristics on disruptive behaviour (Table **3.18**; Figure **3.17**). These studies included data on 34,524 individuals from 6 cohorts from 3 countries. The types of neighbourhood characteristics considered included: density (*k* = 1), disadvantage (*k* = 1), disorganisation (*k* = 1), employment levels (*k* = 1), noise levels (*k* = 1), social cohesion and informal social control (*k* = 1) and violence (*k* = 1).

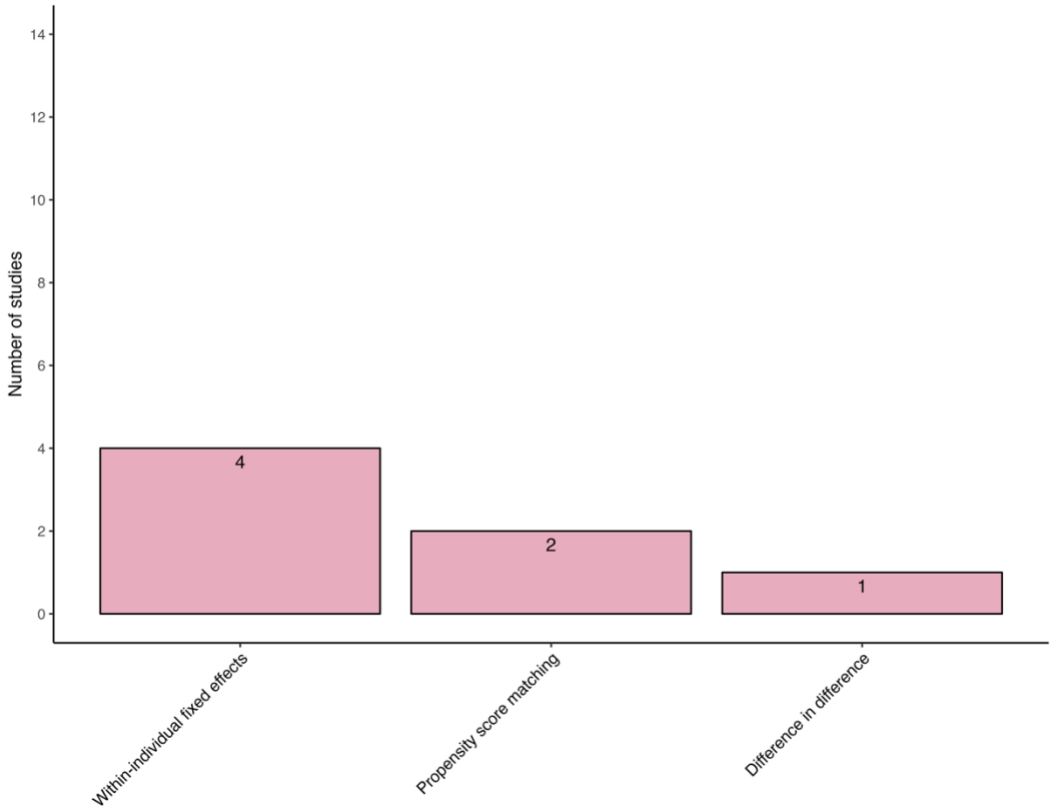


Figure 3.17 *The causal inference methods used in the studies that investigated the influence of neighbourhood characteristics on disruptive behaviour disorder symptoms.*

Table 3.18 *Selected characteristics of studies using causal inference methods to investigate the influence of neighbourhood characteristics on disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Bubonya *et al.,* (2019)	LSAC	Australia	DiD	Employment	EXT	4,089
Harden *et al.,* (2009)	NLSY79	USA	FE	Density	CP	9,440
Humphrey *et al.,* (2017)	ECLS-K	USA	PSM	Disadvantage	EXT	14,960
Ichikawa *et al.,* (2017)	J-SHINE	Japan	FE	Social cohesion and informal social control	EXT	918
Ma *et al.,* (2018)	FFCWS	USA	FE	Dis-organisation	Other	2,472
Ma *et al.,* (2020)	FFCWS	USA	FE	Violence	EXT	2,472
Rudolph *et al.,* (2019)	NCS-A	USA	PSM	Noise	Other	2,645

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: DiD = difference in difference study, PSM = propensity score matching, FE = within-individual fixed effects. Outcomes: CP = conduct problems, EXT = externalising symptoms, Other = other disruptive behaviour disorder.

3.3.5.2 Peer relationships

Three studies used the discordant twin (k = 2) and discordant sibling study designs (k = 1) to investigate the impact of peer relationships on DBDs, including exposure to bullying (k = 1), peer affiliation (k = 1) and having delinquent peers (k = 1; Figure **3.18**; Table **3.19**). The studies included data on 13,775 individuals from 3 cohorts in 2 countries.

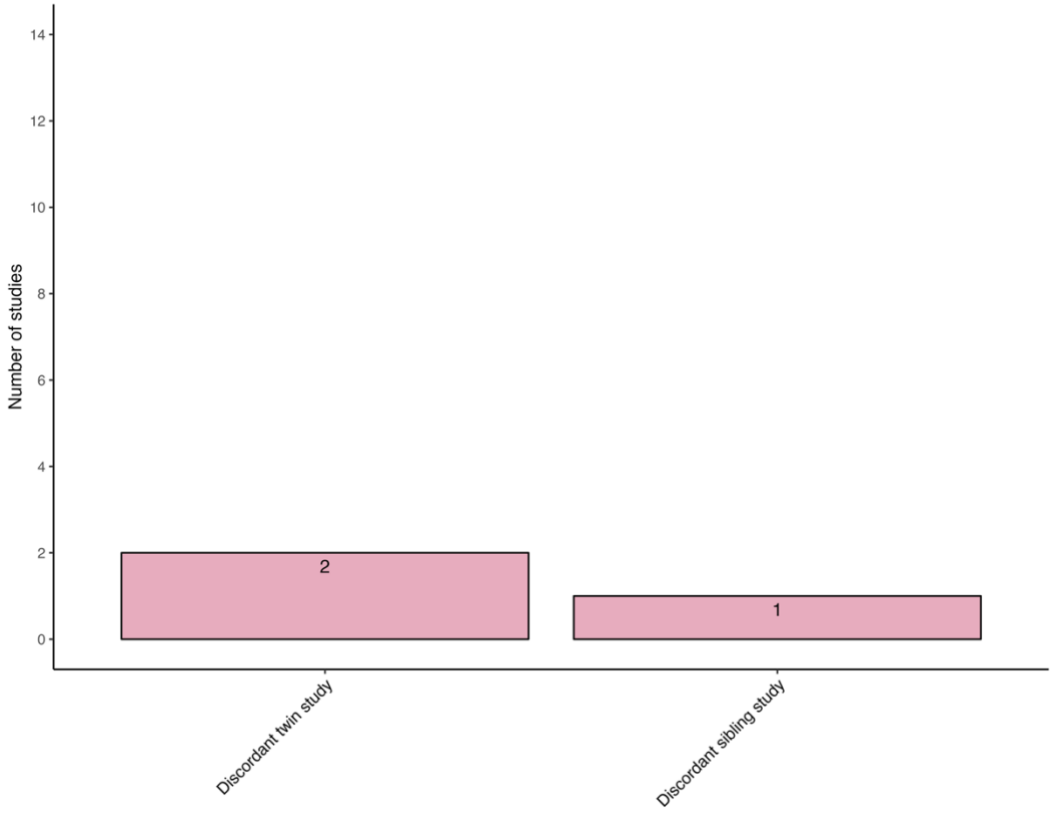


Figure 3.18 *The causal inference methods used in the studies that investigated the influence of peer relationships on disruptive behaviour disorder symptoms.*

Table 3.19 *Selected characteristics of studies using causal inference methods to investigate the influence of peer relationships on disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	Method	Risk factor	Outcome	*n*
Boisvert *et al.,* (2008)	PSID	USA	Sib	Delinquency	EXT	1,759
Burt *et al.,* (2009)	MTFS	USA	Twin	Affiliation	EXT, CD or ODD	908
Singham *et al.,* (2017)	TEDS	UK	Twin	Bullying	CP	11,108

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Sib = discordant sibling study, Twin = discordant twin study. Outcomes: CP = conduct problems, CD = conduct disorder, EXT = externalising symptoms, ODD = oppositional defiant disorder.

3.4 Discussion

I identified 167 studies published since 1980 that examined the impact of 23 putative risk factors on DBD symptoms, using 13 different causal inference methods. These studies included nearly 1 million unique participants across 20 different countries. Most of the research focused on three risk factors: parenting practices (*k* = 39; 21.4%), prenatal exposure to toxins (*k* = 26; 14.3%) and parental internalising symptoms (*k* = 17; 9.3%). Most of the studies used genetically informed family-based methods, such as discordant sibling (*k* = 31; 18.8%), adoption (*k* = 29; 17.1%), discordant twin (*k* = 24; 14.1%) and children of twins (*k* = 12; 7.1%) study designs, but a large proportion also used propensity score matching analyses (*k* = 35; 20.6%) and within-individual fixed effects (*k* = 16; 9.4%).

I identified six risk factors that have been examined by at least ten studies using causal inference methods; these include parenting practices (*k* = 39; 23.4%), prenatal exposure to toxins (*k* = 26; 15.6%), parental internalising symptoms (*k* = 17; 10.2%), adverse childhood experiences (ACEs; *k* = 14; 8.4%), education (*k* = 14; 8.4%), negative family dynamics (*k* = 13; 7.8%) and parental substance use (*k* = 10; 6.0%). Some of these risk factors have already been included in previous reviews of causal inference studies (e.g. ACEs, Baldwin et al., 2023; prenatal exposure to smoking, alcohol, and caffeine, Haan et al., 2022), but it would be useful to triangulate evidence on the other identified risk factors in further meta-analyses to determine whether they are causal.

In comparison, the majority of the 23 risk factors had been examined by fewer than ten studies. These include neighbourhood characteristics (*k* = 7), parental externalising symptoms (*k* = 7), socioeconomic position (*k* = 6), parental education and employment (*k* = 5), birth weight (*k* = 4), breastfeeding (*k* = 3), own substance use (*k* = 3), parental ACEs (*k* = 3), peer relationships (*k* = 3), maternal age at birth (*k* = 2),

adolescent childbearing (k = 1), gaming (k = 1), obstetric complications (k = 1), residential mobility (k = 1), TV viewing (k = 1) and temperament (k = 1). Several putative risk factors, including hormonal and psychophysiological factors, were not included in any study. Future research using causal inference methods should focus on these risk factors to assess whether they are causally related to DBDs.

Some causal inference methods have been underutilised but could be fruitful avenues for future research. The IVF study design was only used in five studies, which examined the effects of breastfeeding, parental externalising symptoms, parental internalising symptoms, parenting practices and prenatal exposure to toxins. Instrumental variable (IV) analyses have been used in five studies to investigate the impact of education and prenatal exposure to toxins, including alcohol and lead. The IVs used for education were random-assignment of childcare programmes or subsidies (Crosby et al., 2010; Herbst & Tekin, 2016) or indicators of preschool (Monnet, 2019). The IVs for prenatal exposure to toxins included maternal drinking during the three months before pregnancy (IV for alcohol exposure; Lund et al., 2019) and melting plant locations (IV for lead exposure; Sampson & Winter, 2018). Four studies used natural experiments to examine socioeconomic position (receipt of an income supplement; Costello et al., 2003, 2010), prenatal exposure to toxins (wartime famine; Neugebauer et al., 1999) and ACEs (institutional abuse; Sonuga-Barke et al., 2017). Two studies have used MR analyses to examine the impact of alcohol use (Chao et al., 2017) and prenatal exposure to alcohol (Murray, Burgess, et al., 2016) on DBDs. One study used the difference-in-difference design to examine the effect of neighbourhood unemployment rates after the Great Recession (Bubonya et al., 2019). Finally, one study used the regression discontinuity design to investigate the effect of education on DBDs, using date of birth as a proxy for school starting age (Dee & Sievertsen, 2018). Although a few combinations of risk factors and causal inference methods are not feasible (e.g. MR analyses on TV viewing), the current review

indicates that there are a lot of potential avenues for future research using causal inference methods to fill the gaps identified here (Figure **3.2**).

3.5 Choice of risk factors and methods

The current review directly informed the choice of risk factors and methods used in this thesis. Firstly, I chose to investigate the impact of parenting practices in my meta-analysis of causal inference methods (Chapter 4) as it was the most studied risk factor, and the studies included a good spread of causal inference methods, which will further triangulation. Secondly, I selected a risk factor, resting heart rate, posited to have a causal relationship with DBDs in previous research but has not yet been investigated using causal inference methods to examine whether this result can be replicated using alternative methods (Chapter 5). Thirdly, I chose a method that has yet to be used to investigate DBDs, g-computation, to examine the impact of early life adversity (Chapter 6).

Key points

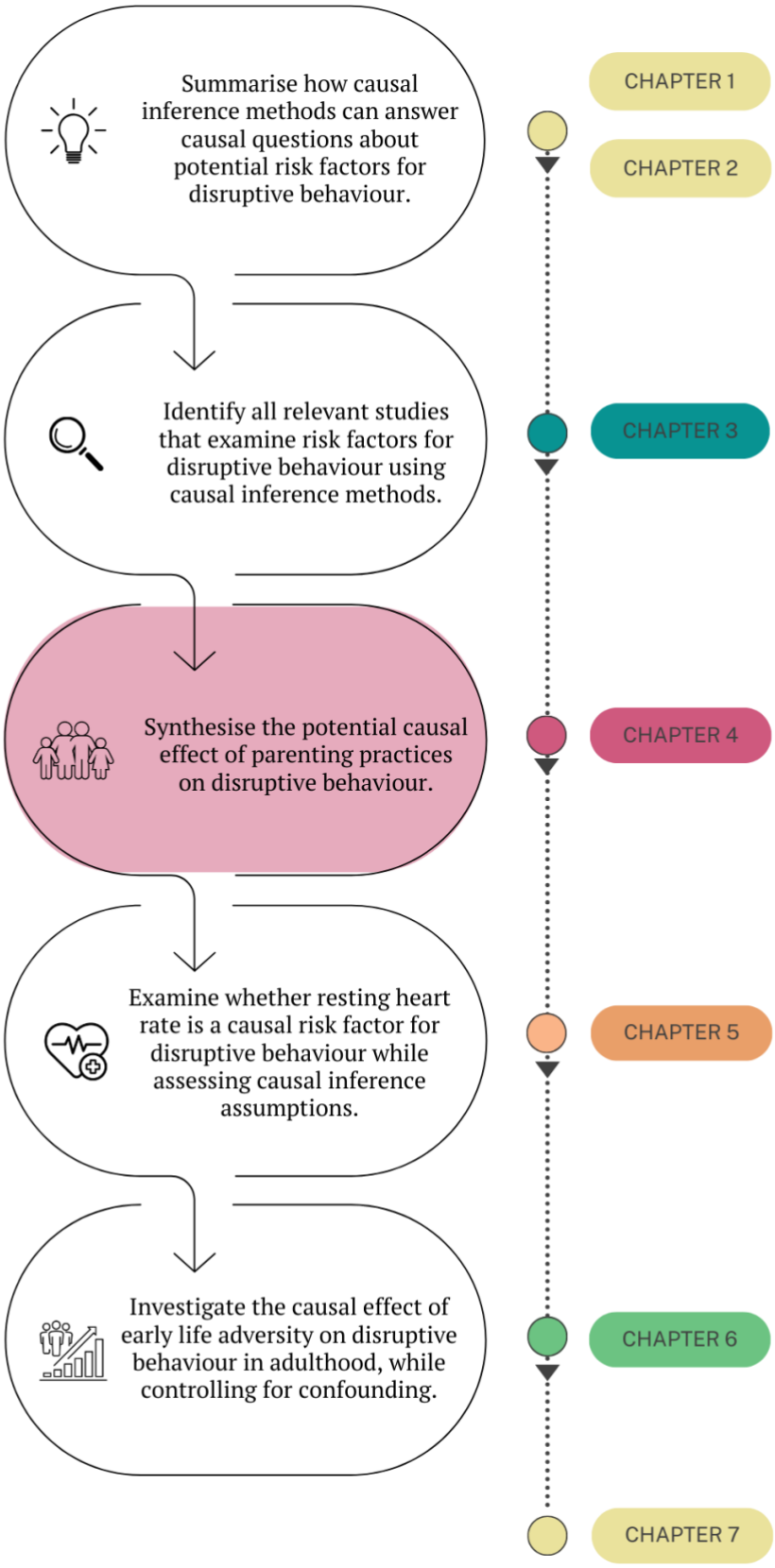
1. I identified 167 studies published since 1980, which included nearly 1 million participants and examined the impact of 23 putative risk factors on DBDs using 13 different causal inference methods.

2. The results highlighted the risk factors frequently examined using causal inference methods (parenting practices, prenatal exposure to toxins and parental internalising symptoms) and the methods commonly employed (discordant sibling studies, adoption studies and propensity score matching analyses).

3. The findings also showed a lack of causal inference studies for most identified risk factors. Furthermore, the results identified several causal inference methods that had rarely (or never) been used to examine risk factors for DBDs.

4. The current review directly informed the choice of risk factors and methods used in this thesis. I chose to quantitatively synthesise evidence for parenting practices from studies using causal inference methods. I also decided to conduct a Mendelian randomisation study of resting heart rate. Finally, I chose to use G-methods to examine early life adversity.

THESIS STRUCTURE



4 POSITIVE AND NEGATIVE PARENTING PRACTICES AND OFFSPRING DISRUPTIVE BEHAVIOUR: A META-ANALYSIS OF STUDIES USING CAUSAL INFERENCE METHODS

Chapter overview

In Chapter 3, I described the results from my systematic review of evidence for all risk factors for DBDs using causal inference methods. In this Chapter, to aid triangulation, I will focus on one risk factor, parenting practices, which was examined in the largest number of studies using the widest variety of causal inference methods. Furthermore, parenting practices are also a crucial element in preventative interventions for DBDs. Therefore, an interesting question arises: Do observational studies on parenting practices report a causal effect on offspring DBDs? If so, do these results vary by the type of causal inference method used, and what can this tell us about other factors influencing the relationship between parenting practices and DBDs? To answer these questions, I estimated the potential causal effect of parenting practices on offspring DBDs in a meta-analysis of studies using causal inference methods.

Publication status: This Chapter is currently under review in *Psychological Bulletin*. Karwatowska, L., Solmi, F., Baldwin, J. R., Jaffee, S., Viding, E., Pingault., J.-B., & De Stavola, B. L. Positive and negative parenting practices, and offspring disruptive behaviour: a meta-analysis of quasi-experimental evidence. (submitted July 2023). *Under review in Psychological Bulletin*.

4.1 Background

As discussed in Chapter 1, the development of disruptive behaviour is influenced by both genetic and environmental risk factors (Fairchild et al., 2019), with many of the latter providing opportunities for intervention. Caregivers play a key role in child development, and the most extensively studied and widely used intervention for disruptive behaviour disorders (DBDs), Parent Management Training (PMT), specifically targets parenting practices (Eyberg et al., 2008; Pilling et al., 2013). It is thought that parenting influences symptoms of DBDs in their offspring through social learning (i.e., offspring learning disruptive behaviours by observing and imitating their parents; Bandura & Walters, 1963) and operant conditioning (i.e., parents modifying their offspring's disruptive behaviour using reward and punishment; Skinner, 1950). PMT aims to improve the parent-child relationship by reducing negative parenting practices, such as harsh, coercive parenting, and promoting more positive behaviours, such as warm, supportive parenting, as well as encouraging involvement and communication between parents and their offspring. Randomised control trials (RCTs) indicate a moderate but highly heterogeneous effect of PMT on child disruptive behaviour (Cohen's d = -0.21 to -0.69; Leijten et al., 2019), with estimates suggesting up to half a standard deviation (SD) reduction in DBD scores.

Although RCTs are considered the "gold standard" for causal inference, they do not typically test which aspect of parenting leads to reduced DBD symptoms, as the interventions usually include a mixture of policies (Leijten et al., 2019). Indeed, PMT targets a suite of positive and negative parenting practices, some, all, or none of which could be the effective mechanism of intervention (Leijten et al., 2022). In contrast to RCTs, studies that use causal inference methods focus on the associations between specific parenting practices (e.g. harsh, coercive parenting) and disruptive behaviours. Therefore, RCTs and observational studies using causal inference methods ask separate but complementary questions.

As discussed in Chapter 1, causal inference methods can help address some of the limitations associated with RCTs. For example, RCTs are expensive to run and, therefore, often have a short follow-up duration, difficulty recruiting a broad range of individuals, and sometimes different levels of attrition by intervention arm, resulting in low external validity (Bärnighausen, Røttingen, et al., 2017; Hernán & Robins, 2016). In comparison, studies using causal inference methods have the potential of higher external validity as they often use data that are readily available and rely on large samples over long-term follow-up (e.g., registry data), which are less prone to attrition (Bärnighausen, Tugwell, et al., 2017). Therefore, when data are available, causal inference methods can provide additional information about the aetiology of DBD symptoms, which can, in turn, inform interventions.

Many systematic reviews and (meta-)meta-analyses exist on data from RCTs (Mingebach et al., 2018) and associational studies based on observational data (Pinquart, 2017; Rothbaum & Weisz, 1994). However, only one narrative review exists that synthesises evidence from other causal inference methods on the impact of harsh parental discipline on a common symptom of DBDs, antisocial behaviour (ASB; Jaffee, Strait, et al., 2012). In their review, Jaffee and colleagues concluded that there was evidence that harsh parental discipline had causal effects on ASB. They also reported evidence of reverse causation and familial confounding. This indicates that some of the association between harsh parental discipline is due to gene-environment correlations (r_{GE}), i.e. shared genetics between parents and their offspring influences the association between the exposure and outcome. Furthermore, the presence of reverse causation suggests that as well as harsh parental discipline leading to more offspring ASB, higher levels of ASB, in turn, invoke more harsh parental discipline (i.e., parent- *and* child-driven effects; Jaffee et al., 2012). However, the studies included in the review were not quantitatively synthesised, and since its publication,

there has been a substantial increase in the number of studies using causal inference methods to investigate this topic.

As such, a quantitative summary of the evidence from studies using causal inference methods to investigate the influence of parenting practices on offspring DBD symptoms is both timely and important. In this Chapter, I identify and summarise studies using causal inference methods that examine this relationship and test whether the results from these studies indicate evidence of causal effects. I also examine whether any reported causal effects vary by the following offspring characteristics and/or study features.

Offspring sex

The prevalence of DBDs is higher in boys than girls, which may mean that boys experience a greater number of risk factors or that these risk factors have a greater impact on boys than girls (Moffitt et al., 2001; Polanczyk et al., 2015). Boys may experience different parenting behaviours to girls, and therefore, I examined whether the effects of parenting on disruptive behaviour varied according to offspring sex.

Offspring age

There may be "sensitive periods" during childhood or adolescence where parenting practices particularly influence child development (Scott et al., 2018; Wachs et al., 2014). Previous findings on the modifying effect of age are inconsistent, with some studies indicating larger effects in early childhood, other studies suggesting larger effects in later childhood and adolescence, and yet other studies finding a consistent effect throughout childhood and adolescence (Gardner et al.,

2019; Jeong et al., 2021; Pinquart, 2017; Rodrigues et al., 2021). I explored this further by considering whether the results varied by offspring age at assessment(s).

Type of disruptive behaviour disorder outcome

It has been suggested that parenting practices may have a greater influence on certain types of DBD symptoms, with larger effects observed for broader, as compared to narrower, measures of disruptive behaviour (Pinquart, 2017). Consequently, I explored whether the results differed according to certain DBD symptoms or diagnoses.

Type of causal inference method

As different types of causal inference methods account for different types of confounding (e.g., genetic confounders or non-shared environmental confounders; Goetghebeur et al., 2020; Pingault et al., 2022), I tested whether this impacted the magnitude of the reported effects.

Time between exposure and outcome assessments

The effect of parenting is thought to be stable over time (Mingebach et al., 2018). I assessed whether the time between exposure and outcome assessments moderated the effect of parenting on DBD symptoms.

Informant for the exposure and outcome

Shared method variance is an artefactual covariance between the exposure and outcome if the measures are reported by the same informant (De Los Reyes et al.,

2009). Consequently, I compared the results when the informant reporting the exposure and outcome were the same versus when they were different.

Study quality

As differences in study characteristics can affect study findings (Lipsey & Wilson, 2001), I checked whether the results differed according to the study's risk of bias rating (see Section 4.2.4).

Maternal or paternal parenting

Although much less studied, paternal parenting is thought to have similar effects to maternal parenting on offspring DBD symptoms (Jeong et al., 2016). To add to the literature, I examined whether the impact of parenting on DBD symptoms differed when mothers' versus fathers' parenting practices were considered.

4.2 Methods

The protocol for my systematic review was preregistered with the PROSPERO database [CRD42020169313] and published in the BMJ Open (Karwatowska et al., 2020). An article based on this Chapter is under review in *Psychological Bulletin* and was reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Shamseer et al., 2015; Appendix C, Table 1) and Meta-Analyses of Observational Studies in Epidemiology (MOOSE; Stroup et al., 2000; Appendix C, Table 2) guidelines. More information on the methods used, including eligibility criteria and definitions of key terms, are available in Chapter 2. However, I will briefly summarise these methods below.

4.2.1 Database searches

I systematically searched Embase, PsycINFO, MEDLINE and Web of Science Core Collection for peer-reviewed studies written in English and published from January 1980 to January 2021, which I later updated to January 2022. A list of the search terms (Table 1) and the search techniques used (Table 2) are reported in Appendix A. These included terms for causal inference methods, adapted from Glanville et al. (2017) to include genetically informed designs and DBDs. To identify studies focussing on parenting practices, I limited the search results to studies that included key terms for parenting (e.g. "parent*").

4.2.2 Definitions

Causal inference methods

I broadly defined causal inference methods as those that estimate population-level causal effects using observational data either by *(a)* relying on an instrument (e.g. regression discontinuity, Mendelian randomisation, difference-in-difference

approaches) or *(b)* confounder-control methods (e.g. extensions to regression-based methods, propensity score matching).

Parenting practices

I defined positive parenting practices as warm, sensitive, or child-centred (e.g., responsive to offspring's needs) and negative parenting practices as harsh or insensitive (e.g., corporal punishment, shouting or threatening behaviour). I treated positive and negative parenting practices as separate constructs as they are thought to have unique influences on offspring disruptive behaviour (A. Hipwell et al., 2008; Oliver et al., 2014; Pettit et al., 1997).

Disruptive behaviour disorder symptoms

I defined the outcome either by symptoms (e.g., conduct problems [CP], externalising [EXT] problems) or clinical diagnoses (e.g., conduct disorder [CD], oppositional defiant disorder [ODD], psychopathy, antisocial personality disorder [ASPD] and dissocial personality disorder [DPD]) associated with disruptive behaviour, which I refer to broadly as DBD symptoms.

4.2.3 Screening and data extraction

Two authors (L.K. and F.S.) independently screened the titles and abstracts of all articles retrieved from the searches. Two authors (L.K. and F.S. *or* B.L.D.S.) reviewed the full texts of all potentially eligible studies. After the full-text screen, two authors (L.K. and J.R.B.) independently extracted data from all eligible studies, including information on sample size, confounder adjustment and effect sizes. The original study authors were contacted when this information was missing or incomplete.

When multiple effect sizes were available, the most stringent estimate (i.e., with the greatest degree of control for confounding) was extracted.

4.2.4 Risk of bias

I adapted the Newcastle-Ottawa scale (Wells et al., 2000) to include questions relevant to causal inference methods, including environmental and genetic confounding, reporting bias and temporal ordering of measures (Appendix A). One author (L.K.) coded study quality and any questions were discussed with two team members (B.L.D.S. and J-B.P.). An overall score was derived by summing the scores across all items (highest score = 10), and the 33rd and 66th percentiles were used to categorise the studies into one of three categories: very high-risk (score below 5.5), high-risk (score between 5.5 and 7) or high-quality (score above 7). One author (L.K.) coded the study quality and discussed any questions with two team members (B.L.D.S. and J-B.P.).

4.2.5 Effect size transformation, interpretation, and significance

Most papers measured parenting practices and DBD symptoms on a continuous scale. If the effect parameters were not already standardised (i.e., reported as [Pearson's correlations] *r*), these were transformed into Pearson's correlations using formulae reported in Appendix C (Table 3). Therefore, the results from the meta-analyses represent the association between 1 standard deviation (1 SD) difference in a standardised parenting practices score and a standardised offspring DBD score. If standard errors of the reported parameters were not available, I used the sample sizes and reported *p*-values to calculate them.

4.2.6 Multilevel random effects model

All analyses were conducted in R (4.1.0) using the *metafor (version 4.3-7;* Viechtbauer, 2010) package. As most studies (81%) reported estimates for multiple measures of parenting practices and many studies (54%) used data from the same data sources (i.e. the same cohort), I fitted 3-level linear random-effects models (Assink & Wibbelink, 2016) with the reported effect estimate nested within study nested within cohorts (see Figure **4.1**), which resulted in an overall "pooled" *r*.

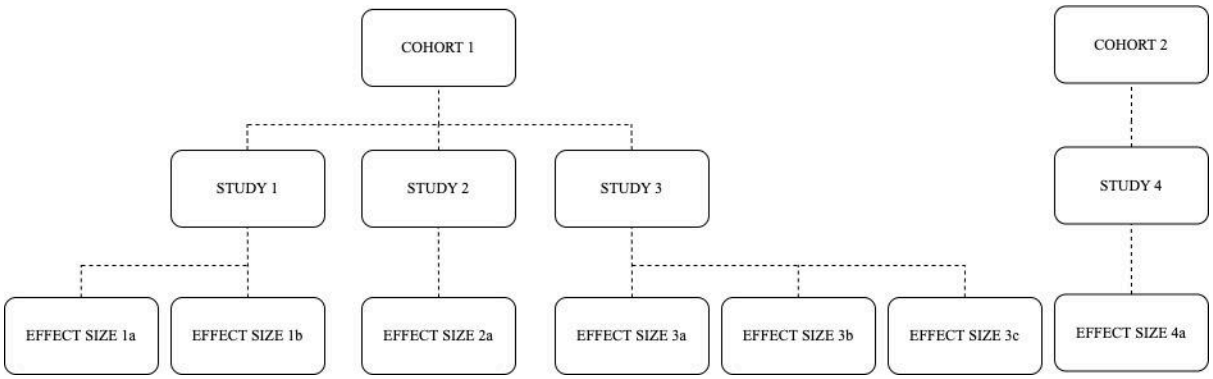


Figure 4.1 *A simplified representation of the data structure (effect sizes nested within studies nested within cohorts) accounted for in the 3-level random-effects model.*

To evaluate publication bias, I created funnel plots to check for asymmetry in the distribution of estimates according to their precision. I also conducted various additional analyses, including Egger's test of heterogeneity (Rodgers & Pustejovsky, 2021) and leave-one-out analyses to recalculate the Egger's test when certain effect estimates were excluded (Viechtbauer & Cheung, 2010).

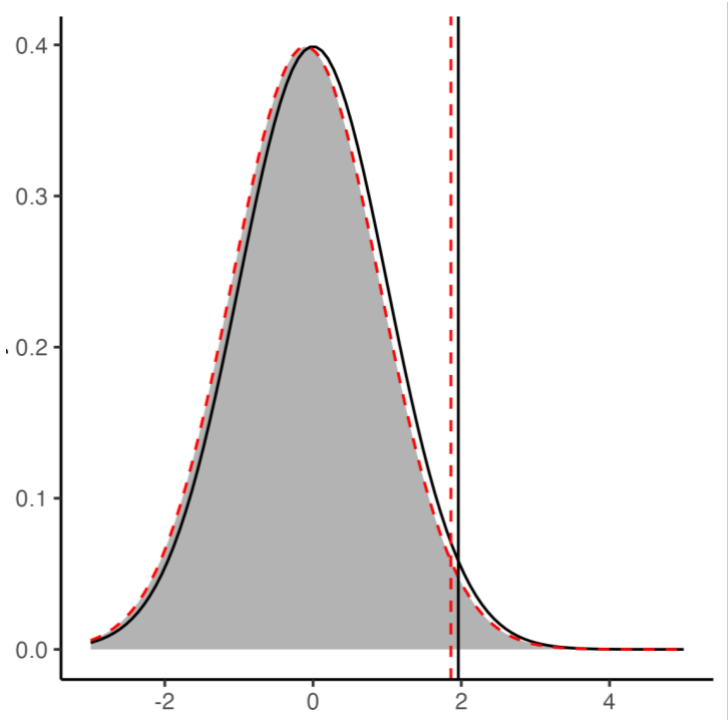
I also examined potential heterogeneity using the *Q* and I^2 statistics. I interpreted an I^2 of more than 50% indicating moderate heterogeneity (Higgins & Thompson, 2002). To further investigate sources of heterogeneity, I conducted another set of leave-one-out analyses where I recalculated the *Q* and I^2 statistics to see if they changed when certain effect estimates were excluded.

4.2.7 Subgroup analyses

I performed subgroup analyses using *a-priori* factors (i.e., published in the pre-registered protocol; Karwatowska et al., 2020) to check whether the pooled estimates differed according to participant characteristics and/or study features. The significance of between-group heterogeneity was assessed by the Wald test (Wald, 1943). Pooled estimates were calculated if there were at least three effect sizes in each category/level and if there was no strong evidence of between-study heterogeneity.

4.2.8 Calculation of the population attributable impact of parenting

I estimated the impact of intervening on parenting, which I call the "population attributable impact" of parenting, as the number of cases of clinical-level DBD symptoms that could be prevented if an effective intervention were available. Assuming the population prevalence of clinically relevant symptoms of DBD is 5.7%, as estimated by Polanczyk et al. (2015), and that this corresponds to the top 2.5% "tail" of the distribution of a standardised DBD score (normally distributed with a mean = 0 and *SD* = 1), I derived the DBD score value above which a diagnosis would be recorded (denoted by 'z'; i.e. the standardised score). I then used the estimated pooled meta-analytic effect (assuming it is causal) to deduce how much the mean DBD score ('z') would change if parenting practices changed in line with the estimated 0.4SD change estimated in universal parenting programmes (Jeong et al., 2021). I then re-calculated the area of the tail that would be greater than our new 'z.' The difference between the two areas can be interpreted as the total number of individuals who would have previously exhibited clinical levels of DBD symptoms but would no longer reach the clinical threshold after the hypothetical intervention. I have visualised the population attributable impact in Figure **4.2**.



Note. the black vertical line represents 'z' the DBD score value above which a diagnosis would be recorded, the red vertical line represents the new value of 'z' after a change in parenting practices.

Figure 4.2 *Population attributable impact of negative parenting.*

4.3 Results

4.3.1 Search results

The study selection procedures are summarised in Figure **4.3**. I identified 41 studies published between 1996 and 2021, which examined data from 27 distinct cohorts. The total analytic sample was 36,661 individuals (48.2% female), the mean age at which parenting practices were assessed was 9.96 years and the mean age at which DBD symptoms were assessed was 11.20 years. Further information on the included studies can be found in Table **4.1** below.

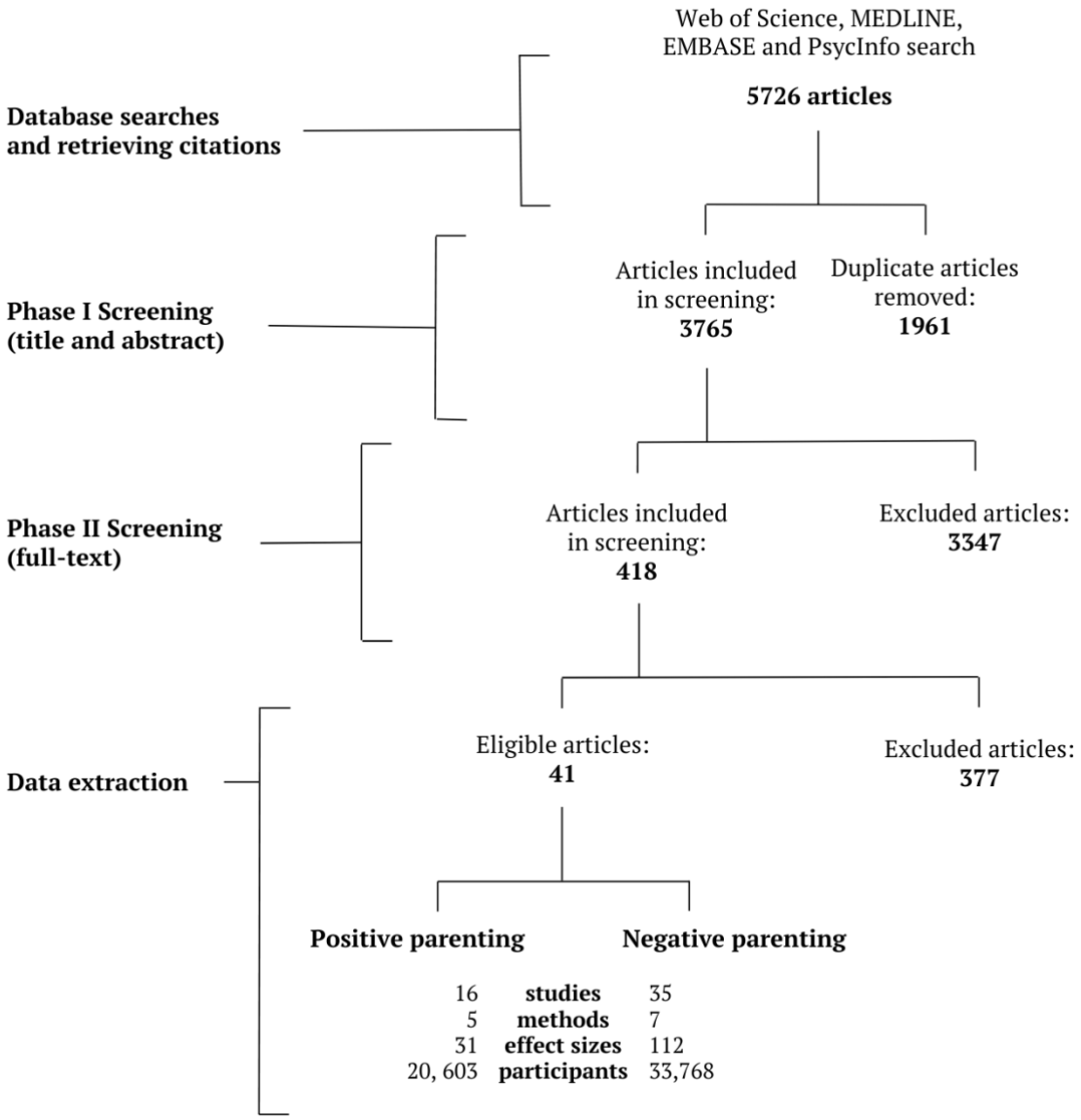


Figure 4.3 *PRISMA flow diagram of search results.*

Table 4.1 *Selected characteristics of studies using causal inference methods to investigate parenting practices and offspring disruptive behaviour disorder symptoms.*

Reference	Cohort	Country	N_{IND}	N_{OBS}	N_{ES}	Method	Sex	Design	RoB	Positive	Negative
Anthony et al., 2019	WACS	UK	62	62	1	Adopt	Mixed	L	7.0	Parental Warmth	-
Asbury et al., 2003	TEDS	UK	4,268	2,134	2	Twin	Mixed	CS	5.0	-	Harsh Discipline; Parental Feeling
Asbury et al., 2006	TEDS	UK	4,090	2,045	3	Twin	Mixed with male and female subsamples	L	7.0	Parent-Child Communication	Parental Feeling; Harsh Discipline
Barnett & Scaramella, 2013	FFCWS	USA	274	137	2	Sib	Mixed with male and female subsamples	L	6.0	Parental Warmth	Coercive Parenting
Besemer et al., 2016	PYS	USA	499	499	3	FE	Males	L	8.0	-	Harsh Discipline; Parental Involvement; Parent-Child Communication
Boisvert & Wright, 2008	PSID	USA	578	289	2	Sib	Mixed with male and female subsamples	L	7.0	Parental Warmth; Parental Monitoring	-
Boyle et al., 2004	OCHS; NLSCY; NLSY79	Canada	7,392	3,696	5	Sib	Mixed	CS	6.0	Parental Involvement; Parental Warmth	Harsh Discipline; Parental Hostility; Physical Discipline
Caspi et al., 2004	E-Risk	UK	1,212	606	3	Twin	Mixed	L	9.5	Expressed Emotion; Parental Warmth	Parental Criticism
Cecil et al., 2012	TEDS	UK	5,184	2,592	2	Twin	Mixed	L	7.0	-	Harsh Discipline; Parental Feeling

Reference	Cohort	Country	N_{IND}	N_{OBS}	N_{ES}	Method	Sex	Design	RoB	Positive	Negative
Cree et al., 2021	EGDS	USA	337	337	1	Adopt	Mixed	L	8.0	-	Overreactive Parenting
Deater-Deckard & Petrill, 2004	N2CAP	USA	224	224	1	Adopt	Mixed	L	5.5	Parent-Child Relationship	-
Ganiban et al., 2021	EGDS	USA	361	361	2	Adopt	Mixed	L	8.0	-	Parental Involvement; Overreactive Parenting
Glover et al., 2010	N2CAP	USA	85	85	2	Adopt	Mixed	CS	5.0	Expressed Emotion	Parental Feeling
Harold et al., 2012	CardiffIVF	UK/USA	207	207	1	IVF	Mixed	CS	5.0	-	Parental Hostility
Harold et al., 2013	EGDS	USA	218	218	1	Adopt	Mixed	CS	7.5	-	Parental Hostility
Hou et al., 2013	BeTwiSt	China	690	345	2	Twin	Mixed	L	9.5	Parental Warmth	Parental Hostility
Klahr, McGue, et al., 2011	SIBS	USA	672	405	1	Adopt	Mixed	L	9.5	-	Parent-Child Conflict
Klahr, Rueter, et al., 2011	SIBS	USA	396	396	2	Adopt	Mixed	CS	6.5	-	Parent-Child Conflict; Coercive Parenting
Latham et al., 2017	TFaB	UK	212	212	1	Sib	Mixed	L	7.5	-	Coercive Parenting
Lipscomb et al., 2014	EGDS	USA	233	233	1	Adopt	Mixed	L	7.0	-	Overreactive Parenting
Long et al., 2015	VATSPSUD	USA	2,606	1,303	3	Twin	Mixed	CS	5.5	-	Low Parental Warmth; Overprotective Parenting; Harsh Discipline
Lysenko et al., 2013	TEDS	UK	9,096	4,548	1	Sib	Separate male and female samples	L	6.5	-	Harsh Discipline
Marceau et al., 2013	EGDS	USA	561	561	1	Adopt	Mixed	L	7.0	-	Overreactive Parenting

Reference	Cohort	Country	N_{IND}	N_{OBS}	N_{ES}	Method	Sex	Design	RoB	Positive	Negative
Mark & Pike, 2017	SBS	UK	156	78	2	Sib	Mixed	L	5.5	Parent-Child Relationship	Parent-Child Conflict
Meunier et al., 2012	HBHC	Canada	809	599	1	Sib	Mixed	CS	6.0	Parent-Child Relationship	-
Morcillo et al., 2011	BYS	USA/Puerto Rico	653	653	1	PSM	Separate male and female samples	L	9.0	Family Bonding	-
Narusyte et al., 2011	TOSS	Sweden	3,540	3,540	1	CoT	Mixed	L	7.0	-	Parental Criticism
Oliver, 2015	TEDS	UK	6,308	3,154	1	Twin	Mixed	L	7.0	-	Parental Feeling
Paine et al., 2021	WACS	UK	96	96	1	Adopt	Mixed	L	8.5	Parental Warmth	-
Pike et al., 1996	NEAD	USA	186	93	1	Twin	Mixed	CS	5.0	-	Parent-Child Conflict
Reuben et al., 2016	EGDS	USA	225	225	2	Adopt	Mixed	L	7.5	Parental Warmth	Overreactive Parenting
Richmond & Stocker, 2006	Not reported	Not reported	186	93	1	Sib	Mixed	L	4.5	-	Parental Hostility
Richmond & Stocker, 2009	Not reported	Not reported	228	114	1	Sib	Mixed	L	4.5	-	Parental Hostility
Riggins-Caspers et al., 2003	Not reported	Not reported	150	150	2	Adopt	Mixed	L	5.0	-	Physical Discipline; Harsh Discipline
Rolon-Arroyo et al., 2018	Not reported	USA	162	162	1	FE	Mixed	L	6.5	-	Overreactive Parenting
Roos et al., 2016	EGDS	USA	293	293	1	Adopt	Mixed	L	8.5	-	Parental Involvement
Samek et al., 2014	SIBS	USA	533	533	3	Adopt	Mixed	L	7.0	-	Parental Involvement; Parent-Child Conflict; Parent-Child Relationship
Shelton et al., 2008	CaStANET	UK	462	217	2	Twin	Mixed	L	7.0	-	Parental Hostility; Parental Warmth

Reference	Cohort	Country	N_{IND}	N_{OBS}	N_{ES}	Method	Sex	Design	RoB	Positive	Negative
Shewark et al., 2021	EGDS	USA	561	561	2	Adopt	Mixed	L	7.0	High Parental Warmth	Parental Hostility
Viding et al., 2009	TEDS	UK	4,056	2,028	1	Twin	Mixed	CS	7.5	-	Harsh Discipline
Waller et al., 2018	TBED-C	USA	374	187	2	Twin	Mixed	CS	5.5	Parental Warmth	Harsh Discipline

Abbreviations. Cohort abbreviations can be found in the List of Abbreviations. Causal inference methods: Adopt = adoption study, CoT = children of twins study, DiD = difference in difference study, Sib = discordant sibling study, Twin = discordant twin study, RG = fuzzy regression discontinuity design, IV = instrumental variable analyses, IPTW = inverse probability of treatment weighting, IVF = in-vitro fertilisation study, MR = Mendelian randomisation, NE = natural experiment, PSM = propensity score matching, FE = within-individual fixed effects. Design: L = longitudinal, CS = cross-sectional.

4.3.2 Descriptive analyses

The 41 studies included in the meta-analysis assessed eight types of positive parenting practices and 17 types of negative parenting practices (Table **4.2**). The studies used seven different causal inference methods, including the adoption design, discordant twin (including both MZ and DZ twins) and discordant sibling designs, within-person fixed effects analyses, extended children of twins' design, in-vitro fertilisation (IVF) design, and propensity score matching analyses (Table **4.3**).

Table 4.2 *Descriptive summary of the positive and negative parenting practices measured in the studies included in the meta-analysis.*

Measure	*k (%)*	*ES (%)*
Positive parenting measures		
Parental Warmth	9 (47.4%)	12 (38.7%)
Parent-Child Relationship	3 (15.8%)	4 (12.9%)
Expressed Emotion	2 (10.5%)	4 (12.9%)
Family Bonding	1 (5.3%)	4 (12.9%)
High Parental Warmth	1 (5.3%)	2 (6.5%)
Parent-Child Communication	1 (5.3%)	2 (6.5%)
Parental Involvement	1 (5.3%)	2 (6.5%)
Parental Monitoring	1 (5.3%)	1 (3.2%)
Negative parenting measures		
Harsh Discipline	10 (20%)	21 (18.8%)
Parental Hostility	8 (16%)	18 (16.1%)
Overreactive Parenting	6 (12%)	8 (7.1%)
Parent-Child Conflict	5 (10%)	29 (25.9%)
Parental Feeling	5 (10%)	6 (5.4%)
Parental Involvement	4 (8%)	4 (3.6%)
Coercive Parenting	3 (6%)	5 (4.5%)
Parental Criticism	2 (4%)	6 (5.4%)
Physical Discipline	2 (4%)	3 (2.7%)
Low Parental Warmth	1 (2%)	4 (3.6%)
Overprotective Parenting	1 (2%)	4 (3.6%)
Parent-Child Communication	1 (2%)	1 (0.9%)
Parent-Child Relationship	1 (2%)	1 (0.9%)
Parental Warmth	1 (2%)	2 (1.8%)

Note. some studies report estimates for both positive and negative parenting. Abbreviations: k = number of studies, ES = number of effect sizes, % = percentage

Table **4.3** includes further descriptive information. From the 41 studies, I extracted 143 adjusted effect sizes (*n* [individuals] = 36,661), including 31 effect sizes for positive parenting measures (*n* = 20,603) and 112 effect sizes for negative parenting measures (*n* = 33,768). Each effect size is reported in Appendix C, separately for positive (Table 4) and negative parenting measures (Table 5). Across all studies, most of the measures concerned maternal parenting practices, with no studies including only measures of paternal parenting practices and 12 studies (29.3%) including measures of maternal parenting practices only. Of the studies that looked at maternal and paternal behaviour together, 13 studies (31.7%) included separate measures of maternal and paternal behaviour, and 16 studies (39.0%) included combined measures of maternal and paternal parenting practices. Female caregivers far outnumbered male caregivers, with the average number of mothers in the sample being 78.7% versus 20.4% for the average number of fathers.

Regarding offspring characteristics, most studies included mixed-sex cohorts (*k* [number of studies] = 38; 78%). Two studies included female-only cohorts (4.9%) and three included male-only cohorts (7.3%). Ancestry was reported in just over half of the studies (*k* =17; 56.7%). In twelve cohorts, the majority of ancestry was White (29.3%); in two cohorts, it was Hispanic (4.9%); in two cohorts, it was Asian (4.9%); and in one cohort, it was African American (2.4%). Most studies were longitudinal (*k* = 31; 73.8%), with repeated measures available on participants over time. Over 20% of studies did not account for any covariates. Of the studies that adjusted for at least one covariate, the most common was offspring sex (*k* = 18; 19.4%) and offspring age (*k* = 15; 16.1%).

Table 4.3 *Descriptive summary of the participant characteristics and study features of the studies included in the meta-analysis.*

Characteristic	*k*	%
Average percentage female	-	48.2
Average percentage of mothers	-	78.7
Average percentage of fathers	-	20.4
Majority ancestry †		
White	12	40
Asian	2	6.7
Hispanic	2	6.7
African American	1	3.3
Not reported	13	43.3
Year of publication		
1996 - 2001	1	2.4
2002 - 2006	5	12.2
2007 - 2011	7	17.1
2012 - 2016	16	39.0
2017 - 2021	12	29.3
Geographical region ††		
United States of America	22	51.2
United Kingdom	13	30.2
Canada	2	4.7
China	1	2.3
Sweden	1	2.3
Puerto Rico	1	2.3
Not reported	3	7.0
Cohort ††		
Early Growth and Development Study	8	18.6
Twins Early Development Study	6	14
Sibling Interaction and Behavior Study	3	7.0
Northeast-Northwest Collaborative Adoption Projects	2	4.7
Wales Adoption Cohort Study	2	4.7
National Longitudinal Study of Youth	2	4.7
Beijing Twin Study	1	2.3
Boricua Youth Study	1	2.3
Cardiff IVF study	1	2.3
Cardiff Study of All Wales and North West of England Twins	1	2.3
Environmental Risk Longitudinal Twin Study	1	2.3
Fragile Families and Child Wellbeing Study	1	2.3
Healthy Babies Healthy Children	1	2.3
Nonshared Environment and Adolescent Development project	1	2.3

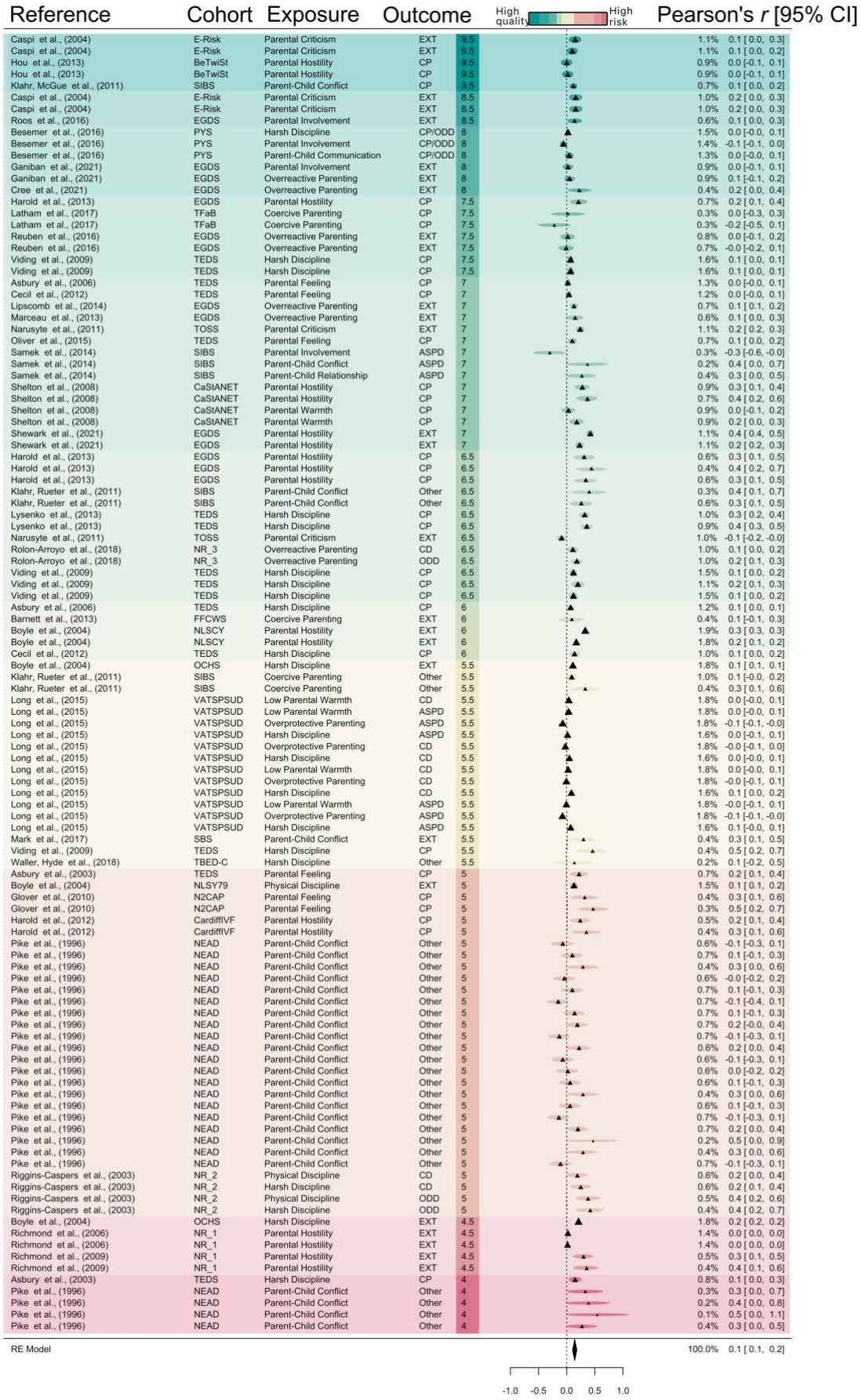
Characteristic	*k*	%
Ontario Child Health Study	1	2.3
Panel Study of Income Dynamics	1	2.3
Pittsburgh Youth Study	1	2.3
Sisters and Brothers Study	1	2.3
The Twins, Family, and Behaviour	1	2.3
Twin Study of Behavioral and Emotional Development in Children	1	2.3
Not reported	4	9.3
Causal inference method ††		
Adoption study	16	39.0
Discordant twin study	11	26.8
Discordant sibling study	9	22.0
Within-person fixed effects	2	4.9
Extended children of twins' study	1	2.4
IVF study	1	2.4
Propensity score matching	1	2.4
Study design ††		
Longitudinal	31	73.8
Cross-sectional	11	26.2
Informants for the exposure and outcome ††		
Discordant	25	53.2
Concordant	22	46.8
Number of covariates in analyses		
0	9	22.0
1	5	12.2
2	13	31.7
3	4	9.8
4	4	9.8
5	3	7.3
6	2	4.9
7	1	2.4
Type of covariates in analyses ††		
Child sex	18	19.4
Child age	15	16.1
Adoption factors	8	8.6
Marital status/quality	7	7.5
Prior DBD	7	7.5
Obstetric complications	6	6.5
Other factors	6	6.5
Socioeconomic factors	5	5.4
Prior parenting	5	5.4

Characteristic	*k*	%
Other parenting factors	3	3.2
Parental psychopathology	3	3.2
In-utero exposure to toxins	3	3.2
Ancestry	2	2.2
Home environment	2	2.2
Interactions between variables	1	1.1
Birth order	1	1.1
Parent age	1	1.1

Note. † = calculated from the total number of cohorts; †† = calculated from the total number of effect sizes. Abbreviations: k = number of studies; % = percentage.

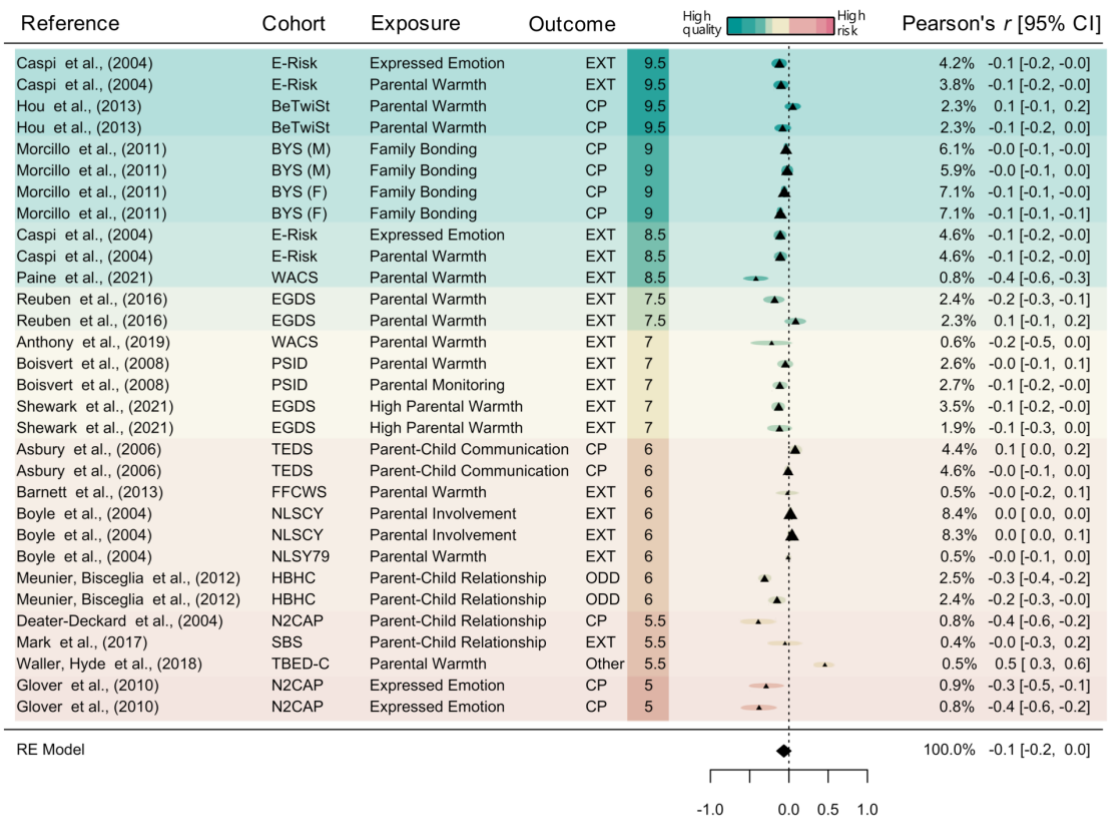
4.3.3 Main meta-analytic results

The multilevel random-effects model for negative parenting found a moderate effect on offspring DBD symptoms ([pooled Pearson's] = 0.142; *95% CI* = 0.104, 0.180; *n* = 33,768). The results suggest that an increase in negative parenting practices is associated with an increase in offspring DBD symptoms. There was little indication of effect heterogeneity (= 22.07%), but as shown in Figure **4.4**, the reported estimates vary by study quality (green to red = high-quality to very high-risk). A descriptive summary of the studies separated by risk of bias category is available in Appendix C (Table 6). The association between negative parenting practices and offspring DBD symptoms was more consistent in the high-quality (i.e., the studies in green) than in the very high-risk studies (i.e., the studies in red). The meta-analysis of positive parenting practices found no association with DBD symptoms and the estimates were more heterogenous (= -0.064; *95% CI* = -0.154, 0.026; *n* = 20,603; = 44.48%; Figure **4.5**).



Note. Results are ordered by risk of bias score, which is also reflected in the colour scheme with green representing studiesthat scored higher on the adapted Newcastle-Ottawa scale ("high-quality") and red representing studies that scored lower ("very high-risk").

Figure 4.4 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of negative parenting practices on offspring disruptive behaviour disorder symptoms.*



Note. Results are ordered by risk of bias score, which is also reflected in the colour scheme with green representing studies that scored higher on the adapted Newcastle-Ottawa scale ("high-quality") and red representing studies that scored lower ("very high-risk").

Figure 4.5 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of positive parenting practices on offspring disruptive behaviour disorder symptoms.*

4.3.4 Sensitivity analyses

4.3.4.1 Publication bias

To assess publication bias, i.e., whether the studies included in the meta-analyses with smaller sample sizes preferentially reported estimates in the expected direction, I visually assessed funnel plots and conducted the Egger's test for asymmetry. Publication bias likely affects the studies in a meta-analysis if the funnel plots are asymmetrical, supported by the *p*-value of the Egger's test below the significance threshold of 0.05. Both checks suggested there was no/limited publication bias for positive parenting measures (*Egger* = 0.533, = 0.47; Figure **4.6** *(i)*). However, the funnel plot for negative parenting was asymmetrical, and the Egger's test was highly

significant, providing strong evidence of publication bias in these studies (*Egger* = 27.38, <0.001; Figure **4.6**, *(ii)*).

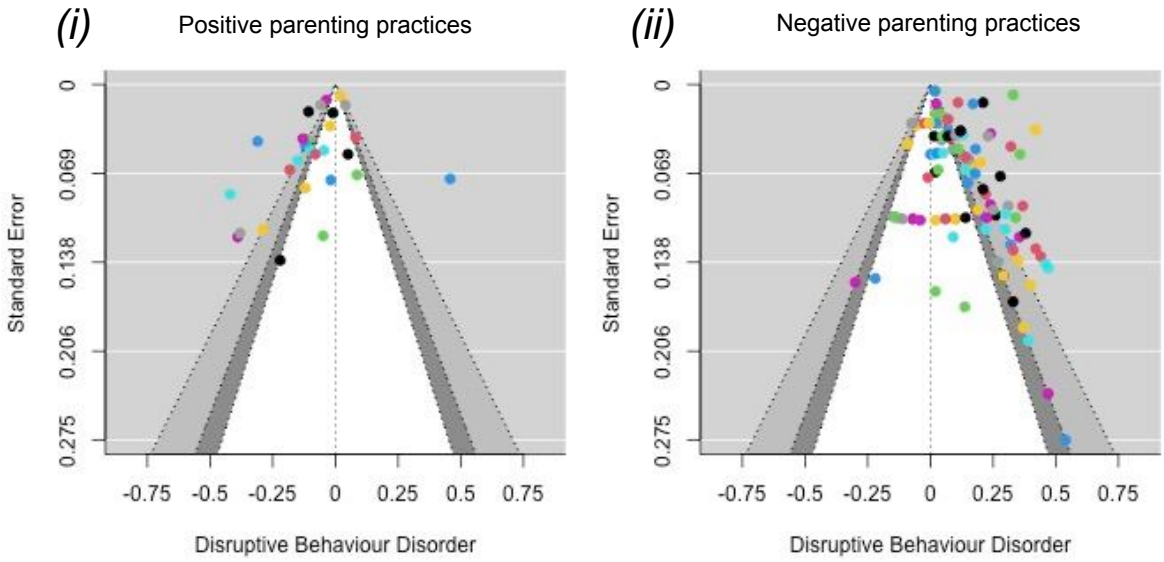


Figure 4.6 *Funnel plots for studies reporting effect estimates for (i) positive parenting (k [number of studies] = 16; ES [number of effect sizes] =31) and (ii) negative parenting (k = 35; ES = 112). The different colours represent different cohorts.*

To understand whether publication bias was driven by any individual studies, effect sizes or study features, I reran the Egger's test for heterogeneity after removing each effect size, study, cohort, type of causal inference method, or risk of bias category from the analyses in turn. I then compared the results from these Egger's tests to the original results to assess whether the *p*-value had changed. For example, if the *p*-value became larger, it would suggest that publication bias was weakened when those selected effect sizes were removed. The resulting *p*-values suggested that no individual effect size, study, or cohort drove the publication bias in the studies reporting on negative parenting. However, when each type of causal inference method was left out of the analyses, the results indicated that publication bias was weakened when discordant twin studies (*k* = 11; ES [number of effect sizes] = 60; Figure **4.7**, *(i)*) and adoption studies (*k* = 13; ES = 28; Figure **4.7**, *(ii)*) were left out in turn.

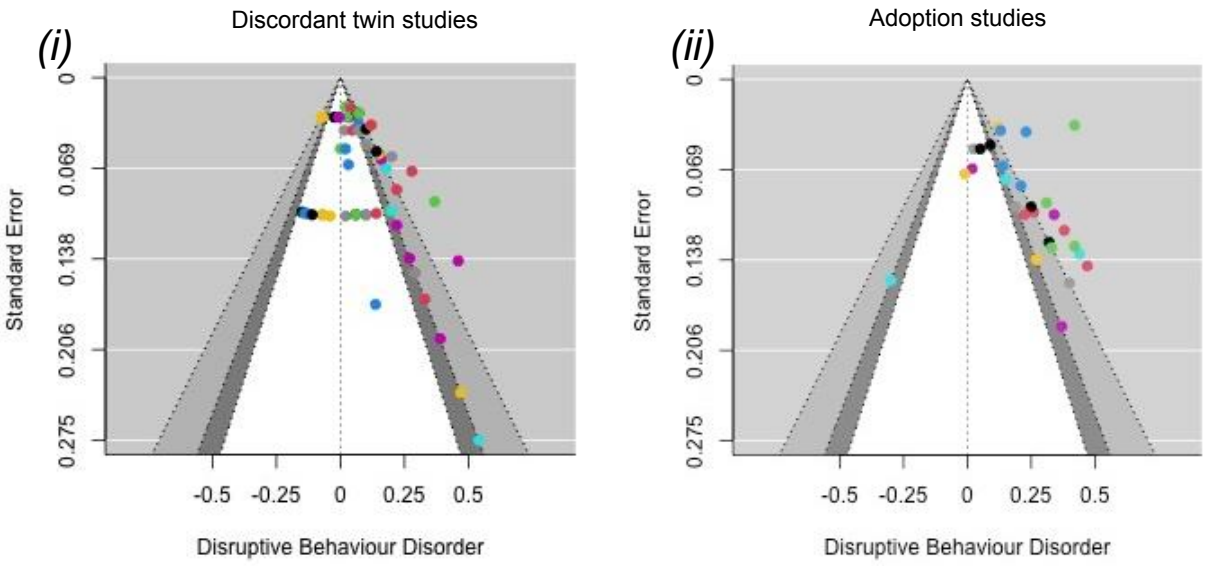


Figure 4.7 *Funnel plots for studies reporting effect estimates for negative parenting practices in (i) discordant twin studies (k [number of studies] = 11; ES [number of effect sizes] = 60) and in (ii) adoption studies (k = 13; ES = 28). The different colours represent different cohorts.*

There was also a gradient to the publication bias depending on the studies' risk of bias, with evidence of publication bias in the studies categorised as very high-risk (*k* = 11; ES = 57; Figure **4.8**, *(i)*) but not those judged as high-risk (*k* = 6; ES = 20; Figure **4.8**, *(ii)*) or high-quality (*k* = 24; ES = 66; Figure **4.8**, *(iii)*).

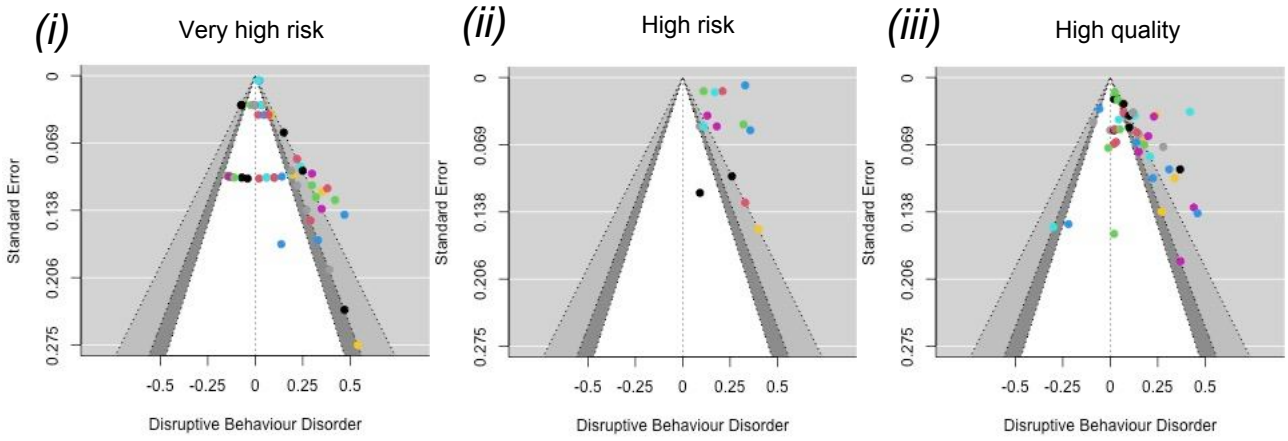


Figure 4.8 *Funnel plots for studies reporting effect estimates for negative parenting practices in studies categorised as (i) very high-risk (k [number of studies] =11; ES [number of effect sizes] = 57), (ii) high-risk (k =6; ES = 20) and (iii) high-quality (k = 24; ES = 66). The different colours represent different cohorts.*

4.3.4.2 Leave-one-out analyses

Leave-one-out analyses indicated that individual effect sizes, studies, or cohorts did not unduly influence the overall pooled estimate for negative parenting practices. After omitting each of the 135 effect sizes, 35 studies, and 20 cohorts in turn, the meta-analytic effect size (*r*) ranged from 0.12 to 0.17 (Appendix C, Figures 1 - 3).

4.3.5 Meta-analytic results from high-quality studies only

As there was evidence of publication bias in very high-risk studies, I re-ran the analyses, removing studies that were judged to be *very* high-risk of bias (= 0.136; *95% CI* = 0.091, 0.180; *n* = 24,763) and then removing those judged to be high-risk of bias, i.e. keeping only effect estimates judged to be high-quality (= 0.104; *95% CI* = 0.053, 0.154; *n* = 16,101). I used the estimate that only included high-quality studies as my most conservative pooled estimate.

4.3.6 Analyses of subgroups

To identify potential sources of heterogeneity in the association between parenting practices and offspring disruptive behaviour, I ran analyses of subgroups defined by a set of pre-specified variables, including participant (e.g. sex, age at outcome) and study features (e.g. type of disruptive behaviour disorder outcome, type of causal inference method used, time between exposure and outcome assessment, whether the exposure and outcomes were reported by the same informant, data quality, maternal vs paternal parenting; see Table **4.4**). The analyses of subgroups were only run for negative parenting measures as the meta-analytic results for positive parenting were not significant, and there were an insufficient number of estimates available for subgroup analyses.

Table 4.4 *Meta-analytic associations between negative parenting practices and offspring disruptive behaviour disorder symptoms for the variables included in the subgroup analyses.*

Term	*k*	*ES*	*n*	*r*	*lCI*	*uCI*
Offspring sex						
Intercept	30	85	33,643	0.148	0.083	0.212
Increasing % Female	30	85	33,643	0.000	-0.001	0.001
Age at outcome assessment						
Intercept	34	108	33,768	0.210	0.140	0.281
Increasing age	34	108	33,768	-0.005	-0.010	0.000
Time between assessments	34	108	33,768	-0.010	-0.026	0.005
Age at exposure assessment						
Intercept	34	108	33,768	0.210	0.140	0.281
Increasing age	34	108	33,768	-0.005	-0.010	0.000
Time between assessments	34	108	33,768	-0.015	-0.031	0.000
Time between assessments						
Intercept	34	108	33,768	0.161	0.112	0.210
Increasing time	34	108	33,768	-0.012	-0.028	0.004
Type of DBD outcome						
Conduct Problems	13	32	11,642	0.159	0.094	0.223
Antisocial Personality Disorder	2	9	3,139	0.047	-0.079	0.172
Conduct Disorder	3	9	2,918	0.080	-0.038	0.197
Externalising Behaviour	14	27	18,367	0.147	0.082	0.211
Oppositional Defiant Disorder	3	6	811	0.152	0.021	0.282
Other DBD	3	29	956	0.135	-0.003	0.274
Type of QE method						
Adoption study	13	28	1,468	0.187	0.130	0.245
Discordant sibling study	7	15	22,362	0.173	0.053	0.294
Discordant twin study	11	60	11,838	0.092	0.039	0.144
Within-person fixed effects	2	5	661	0.067	-0.052	0.186
Informant for exposure and outcome						
Concordant	19	54	28,323	0.191	0.143	0.238
Discordant	22	58	14,453	0.101	0.056	0.145
Data quality						
High quality	24	66	16,101	0.107	0.054	0.160
High risk	6	20	23,133	0.207	0.109	0.304
Very high risk	11	57	8,399	0.170	0.096	0.243
Maternal vs paternal parenting						
Combined	12	24	8,065	0.130	0.062	0.199
Maternal	23	60	32,572	0.146	0.097	0.195
Paternal	12	28	8,278	0.156	0.095	0.216

Abbreviations: k = number of studies; ES = number of effect sizes; n = number of individuals; r = Pearson's r correlation; lCI = lower 95% confidence interval; uCI = upper 95% confidence interval.

4.3.6.1 Participant characteristics

4.3.6.1.1 Offspring age at exposure and outcome assessments

After controlling for the time between exposure and outcome assessments, there was no evidence of an effect of the age of offspring at either exposure (test = 2.31, p = 0.128) or outcome assessment (test = 6.02, = 0.050).

4.3.6.1.2 Offspring sex

The results suggested that the association between negative parenting practices and DBD symptoms did not differ depending on the percentage of females in the sample (test = 0.32, = 0.573).

4.3.6.2 Study features

4.3.6.2.1 Time between exposure and outcome assessment

The time between the exposure and outcome assessments did not influence the association between negative parenting practices and offspring DBD symptoms (test = 2.34, = 0.126).

4.3.6.2.2 Type of disruptive behaviour outcome

The results suggested that the effect of negative parenting practices on offspring disruptive behaviour was similar regardless of the different DBD outcomes, including CP, CD, and ASPD (test = 2.82, = 0.728; see Figure **4.9**).

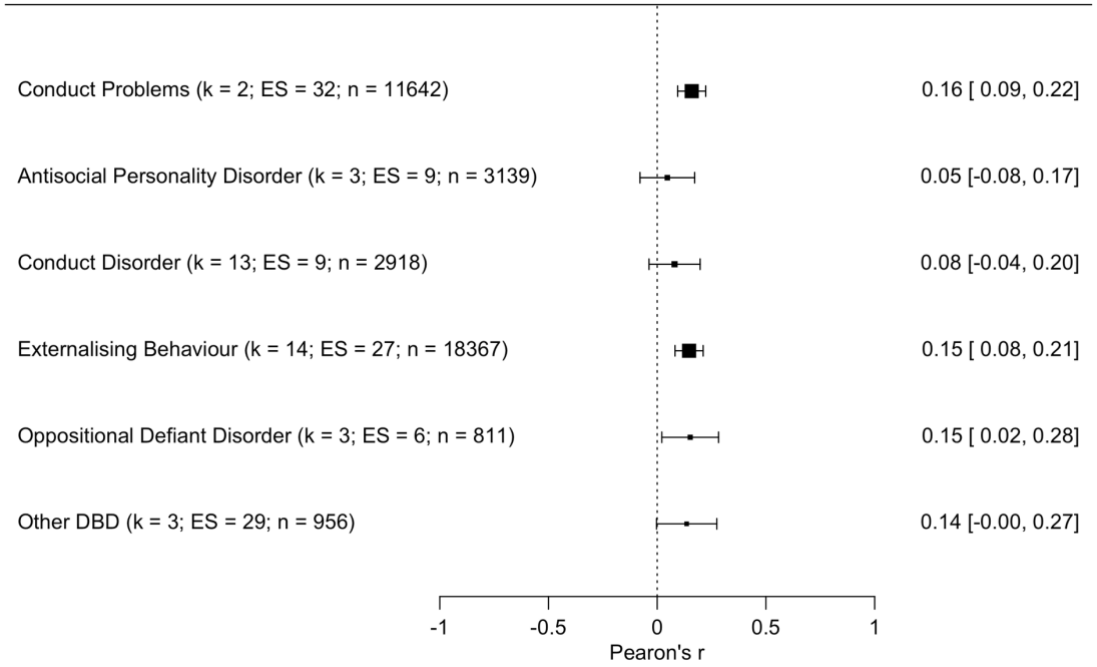


Figure 4.9 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of negative parenting practices on offspring disruptive behaviour disorder symptoms, stratified by type of disruptive behaviour outcome.*

4.3.6.2.3 Type of causal inference method

Although I was not able to include all causal inference methods in the subgroup analyses as there were fewer than three effect sizes, there was some evidence that the magnitude of the effect differed depending on the causal inference method used in the study (test = 4.93, = 0.085). Further analyses suggested that adoption studies (= 0.190, *95% CI* = 0.119, 0.260) reported the largest effects, followed by discordant sibling studies (= 0.173, *95% CI* = 0.053, 0.294), discordant twin studies (*r* = 0.089; *95% CI* = 0.045, 0.133) and finally within-person fixed effect (= 0.07, *95% CI* = -0.040, 0.178; Figure **4.10**).

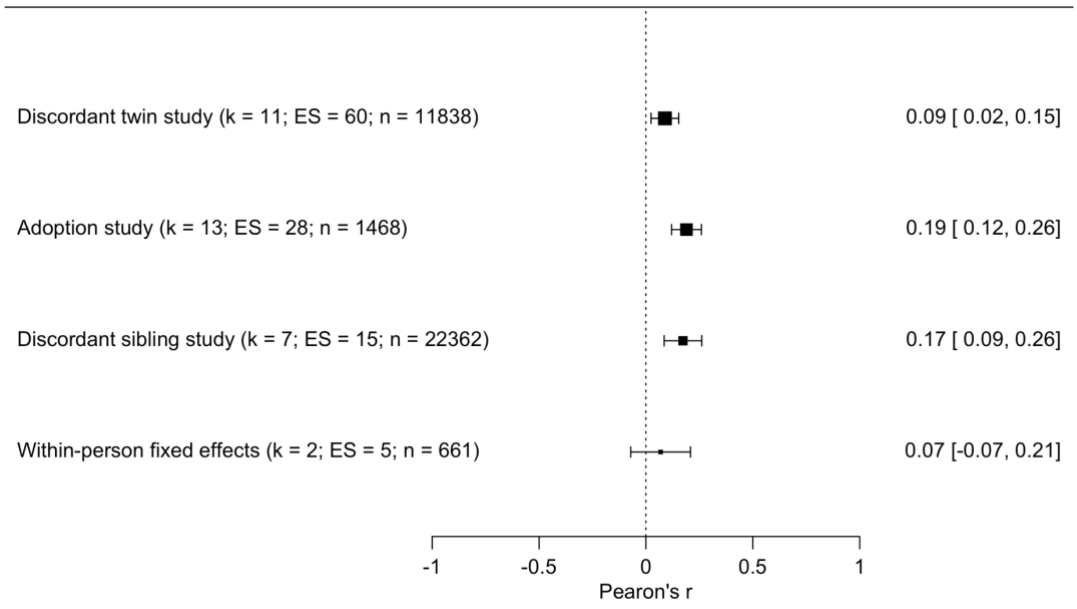


Figure 4.10 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of negative parenting practices on offspring disruptive behaviour disorder symptoms, stratified by type of causal inference method.*

4.3.6.2.4 Informant for the exposure and outcome

There was evidence that the association between negative parenting practices and offspring disruptive behaviour was influenced by whether the exposure and outcome were rated by the same informant (test = 11.582, < 0.001), whereby the association between parenting and disruptive behaviour was reported as smaller when the exposure and outcome were rated by different people (= 0.101, *95% CI* = 0.056, 0.145) compared to when the same informant reported the exposure and outcome (= 0.191, *95% CI* = 0.143, 0.238; Figure **4.11**).

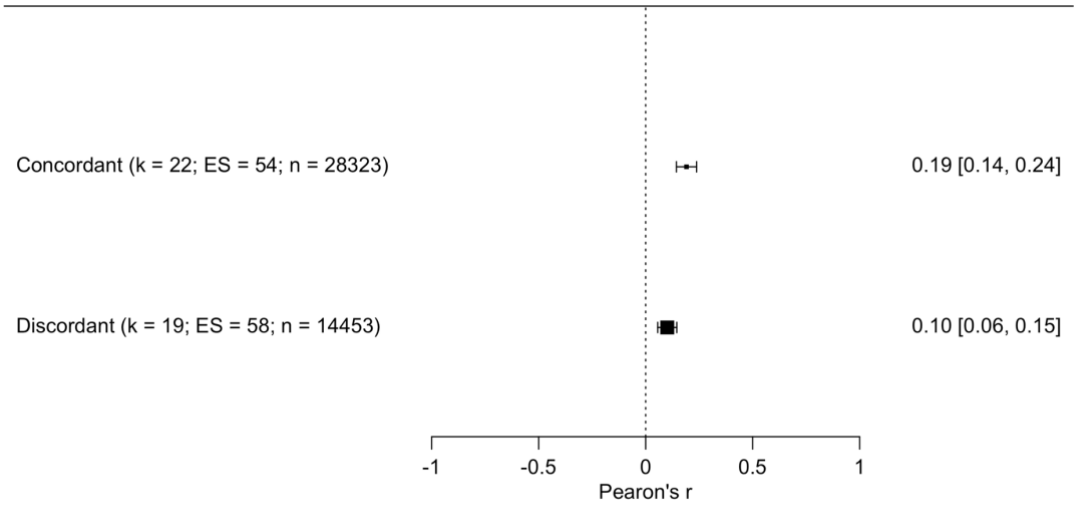


Figure 4.11 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of negative parenting practices on offspring disruptive behaviour disorder symptoms, stratified by the informant for the exposure and outcome.*

4.3.6.2.5 Study quality

The analyses of study quality suggested that the risk of bias of a study was associated with the pooled estimates for negative parenting (= 5.59, = 0.018), with studies judged to be high-risk (= 0.207; *95% CI* = 0.109, 0.304) and very high-risk of bias reporting the largest effects (= 0.170; *95% CI* = 0.096, 0.243), while studies judged to be high-quality reported the smallest effects (= 0.107; *95% CI* = 0.054, 0.160; Figure **4.12**).

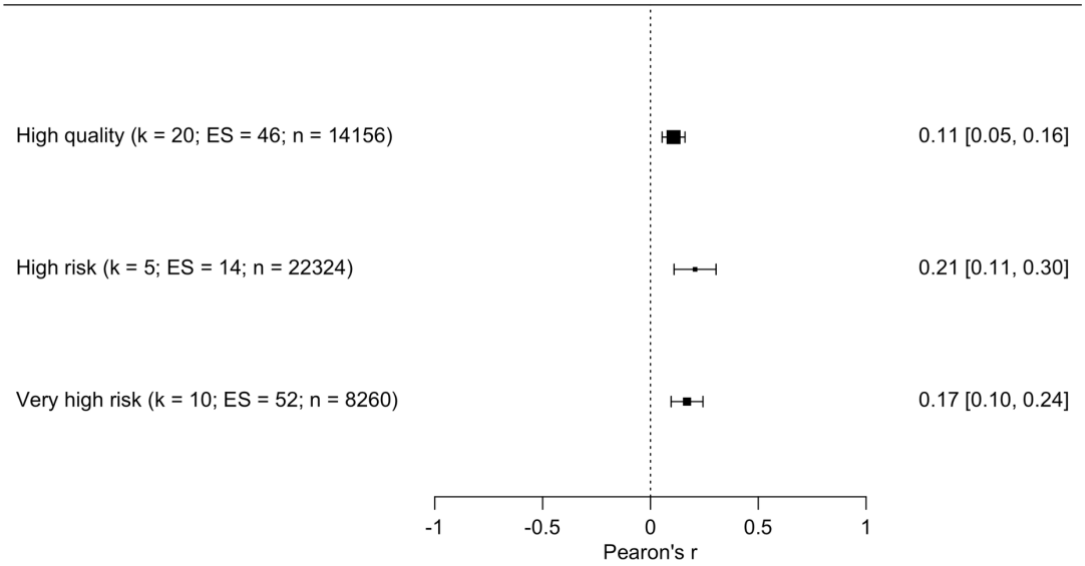


Figure 4.12 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of negative parenting practices on offspring disruptive behaviour disorder symptoms, stratified by the study quality.*

4.3.6.2.6 Maternal versus paternal parenting

The magnitude of the effect did not differ depending on whether the exposure was maternal or paternal parenting practices (test = 0.326, = 0.850; Figure **4.13**).

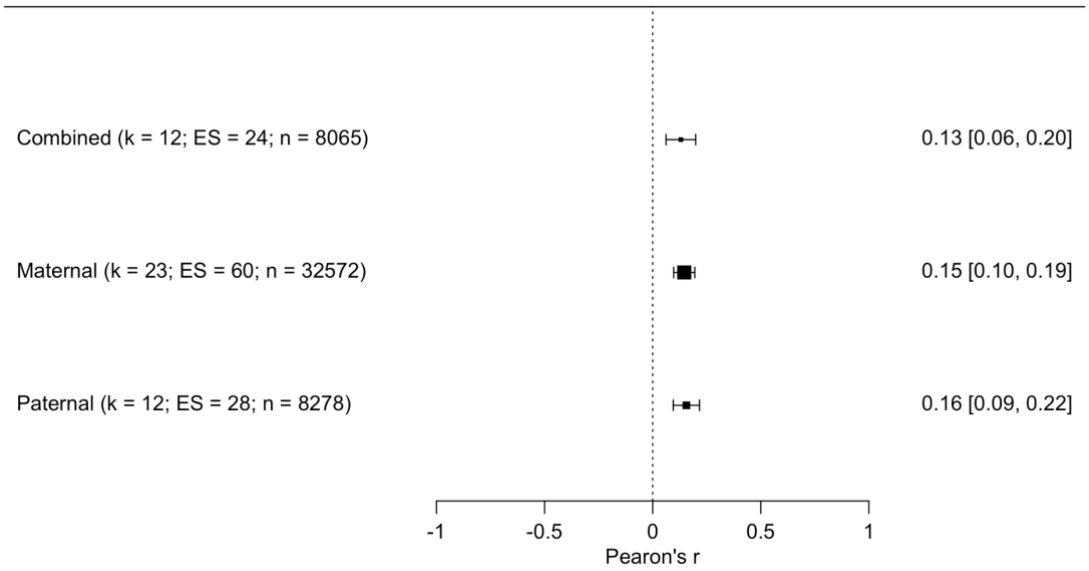


Figure 4.13 *Forest plot of the reported standardised regression coefficients (Pearson's r) for the effect of negative parenting practices on offspring disruptive behaviour disorder symptoms, stratified by maternal or paternal parenting.*

4.3.7 Population attributable impact of negative parenting practices

To estimate the impact of intervening on negative parenting, I calculated the "population attributable impact" of negative parenting, i.e., the number of individuals that might no longer exhibit clinically relevant disruptive behaviour disorder symptoms if the mean score of negative parenting could be reduced. For this calculation, I assumed that the population prevalence of any disruptive behaviour is equal to 5.7% (*n* = 113 million individuals), as previously estimated (Polanczyk et al., 2015), and that a hypothetical intervention changes negative parenting practices by 0.4SD, in line with the changes estimated after universal interventions (Jeong et al., 2021). Using the most conservative pooled meta-analytic estimate (i.e., high-quality studies; *r* = -0.102), I calculated that this hypothetical intervention could lead to a 0.11% reduction in the prevalence of clinically relevant DBD symptoms worldwide, which is the equivalent of 3,614,337 school-aged children worldwide no longer exhibiting clinical levels of DBD symptoms (see Methods 1 in Appendix C for calculation).

4.4 Discussion

This study is the first meta-analysis to synthesise evidence on the effect of parenting practices on offspring disruptive behaviour disorder (DBD) symptoms from studies using causal inference methods. It included 41 studies using data from 27 distinct cohorts with 36,661 independent participants. The findings suggest that negative, but not positive, parenting practices have a small causal effect on offspring DBD symptoms, with the most robust meta-analytic effect of a one unit increase in standardised negative parenting practices score (e.g. equivalent of an increase of 0.58 points on the negativity subscale of the Iowa Family Interaction Rating Scales; Williamson et al., 2011) being associated with 0.102 standard deviation increase in standardised DBD score (e.g. equivalent of an increase of 0.12 points on the Strength and Difficulties Questionnaire; Mieloo et al., 2012). Subgroup analyses suggested that this effect was consistent across offspring sex, age at exposure and outcome assessments, type of DBD outcome, maternal and paternal parenting and time between exposure and outcome assessments. However, it varied by the type of causal inference method used, study quality and whether the same informant rated the exposure and outcome.

The current study builds on the existing literature by triangulating previous meta-analyses of RCTs on parenting interventions (i.e. Parent Management Training; PMT; Mingebach et al., 2018) and associational studies on parenting practices (Cooke et al., 2022). Causal inference methods can, theoretically, address some of the limitations of these other research designs. For example, many associational studies do not adequately adjust for confounding (e.g., genetic or environmental confounders) or address reverse causality (i.e., presence of parent- *and* child-driven effects; Jaffee et al., 2012). In comparison, causal inference methods can account for shared genetics (e.g., family-based studies) and environmental confounders (e.g., baseline characteristics via fixed effects). Furthermore, whereas RCTs often use relatively

small clinical or at-risk samples, studies implementing causal inference methods in observational datasets often use larger samples that are more representative of the general population (Bärnighausen, Tugwell, et al., 2017). In this way, causal inference methods allow us to study normative types of parenting practices and sub-clinical DBD symptoms, which can complement the evidence provided by RCTs.

Along with our main meta-analysis, there were enough studies to examine potential sources of heterogeneity in the effect of negative parenting practices on DBD symptoms. I observed four key findings. First, the results did not differ depending on offspring sex, offspring age, maternal or paternal parenting or type of DBD symptom. Although the prevalence of DBD differs between boys and girls (Polanczyk et al., 2015), our results support previous research indicating that these discrepancies cannot be explained by differences in the effects of negative parenting (Lysenko et al., 2013; Pinquart, 2017). Similarly, the estimated effect of parenting did not vary depending on the age of offspring at either exposure or outcome assessment nor the time between exposure and outcome assessments. Therefore, the derived pooled causal effect of negative parenting is consistent over the age range of the included studies (birth – 37 years). This is consistent with research suggesting that the effects of parenting interventions are similar across a wide range of ages (Gardner et al., 2019). The effect of paternal parenting has been researched less than that of maternal parenting (Jeong et al., 2016, 2021). Therefore, the current review's findings of no difference between maternal and paternal parenting addresses a key gap in the literature. Finally, the effect of parenting practices did not vary according to the different outcomes (e.g., conduct problems [CP], oppositional defiant disorder [ODD], externalising [EXT] symptoms), suggesting that there may be limited utility in specific interventions that target subgroups of DBD symptoms and diagnoses.

Second, our findings suggest that the estimated effect of negative parenting on disruptive behaviour varied according to the type of causal inference method used in the study. Different causal inference methods account for specific types of confounding. A previous systematic review, which included studies using causal inference methods that examined the association between harsh, coercive parenting and antisocial behaviour, highlighted the presence of familial and environmental confounding, which includes shared and non-shared components (Jaffee, Strait, et al., 2012). As discussed in Chapter 1 (Section 1.3.1.1), familial confounding, also called genetic confounding, occurs when the genetic variants shared between parents and children genes influence either the parents' behaviour and/or offspring DBD symptoms (Plomin et al., 1977). Furthermore, shared environmental confounding refers to factors that make family members similar to one another, such as parental psychopathology or family socioeconomic position, which in turn affect the relationship between negative parenting practices and offspring DBDs. On the other hand, non-shared environmental confounding refers to confounding by factors that make family members different to one another, such as relationships with peers.

Interestingly, the studies that reported the largest effects also controlled for the least amount of confounding, with the magnitude of the effects reducing sequentially as the amount of confounding accounted for increased. The largest magnitude effect was reported by adoption designs, followed by discordant sibling designs, discordant twin designs and finally, within-person fixed effects. Firstly, although adoption studies control for genetic confounding, they do not account for environmental confounding, including the prenatal environment, unless specifically controlled for in the analyses (Thapar & Rice, 2021). Secondly, sibling studies do not control for a large amount of genetic and non-shared environmental confounding, including other factors which might lead to changes in parental behaviour (Frisell, 2020; Sjölander et al., 2022; Thapar & Rice, 2021). Thirdly, in addition to controlling for genetic confounding,

discordant twin studies also control for shared environmental confounding but do not control for the non-shared environment (McAdams et al., 2021). Fourthly, fixed-effects studies control for unmeasured time-fixed confounding, both genetic and environmental, as each individual acts as their own control (Gunasekara et al., 2014). However, like the other methods, fixed-effects analyses do not control for time-varying confounding and depend on the frequency of the repeated measures. It should also be noted that only a couple of fixed-effects studies, with only a few effect sizes, were included in the analyses. Therefore, these results should not be overinterpreted.

The subgroup analyses on the type of causal inference method used suggest that some of the effect of negative parenting on DBD symptoms that have been reported in previous associational studies may be confounded by shared genetic or environmental factors. Although causal inference methods theoretically aim to minimise this confounding, our findings suggest that, in practice, no one causal inference method can fully succeed in doing so in isolation (Goetghebeur et al., 2020; Lawlor et al., 2017; Munafò & Smith, 2018). Future research should use innovative and novel causal inference methods, especially designs such as G-methods (Robins, 1986) that have not been implemented yet, to triangulate evidence further.

Third, the results suggest evidence of shared method variance (Podsakoff et al., 2003), whereby estimates are inflated when the same informant reports the exposure and outcome measures, e.g., a parent reporting their parenting behaviour and their offspring's DBD symptoms. Indeed, in the current study, the effects were nearly twice as large when the informants were the same as when they were different. This is consistent with previous meta-analyses on other mental health measures (Francis et al., 2023; Schoeler et al., 2018). Along with more stringent control for confounding, shared method variance may explain why the current results are substantially smaller

than those reported in associational studies (r = 0.22 for parental psychological control; Pinquart, 2017).

Fourth, the higher the study quality, the smaller the reported effects. Indeed, this is linked, in part, to the previous point; the study quality was deemed higher when there was better control for confounders and different informants for the exposure and outcome, including observational measures, which are more immune to recall bias. In addition, higher quality studies were more likely to be longitudinal and control for pre-existing levels of offspring DBDs. This reduces the likelihood of reverse causation (i.e., child-driven effects) whereby children who have more DBD symptoms invoke more negative parenting behaviour (as shown by bi-directional associations in family-based cross-lagged models; Zvara et al., 2018). My most conservative estimate only included high-quality studies; therefore, I am confident that the current findings are not influenced by reverse causation. These results suggest that future research on parenting must account for genetic and environmental confounding and child-driven effects either using study design (e.g., discordant twin study) or analyses (e.g., fixed effects).

Although the current findings cannot be directly compared to those from experimental studies of PMT, it is important to discuss the reasons why the estimates are much lower than those reported in previous meta-analyses of RCTs (Mingebach et al., 2018). Firstly, RCTs not only target parenting practices, but they also influence other factors that may also be risk factors for offspring DBDs, such as parental relationship quality and psychopathology (Jeong et al., 2021; Weber et al., 2019). The studies included in this review examined parenting practices only, and most attempted to control for other variables, such as parental symptoms of depression and marital conflict. Therefore, the findings estimate the causal effect of negative parenting practices *in the absence* of these other factors.

Next, RCTs often use at-risk or clinical samples, and intervention effects are stronger when offspring DBD symptoms are more severe (Menting et al., 2013). For example, in a recent meta-analysis of RCTs of different types of PMT, the magnitude of effects increased as the "level of prevention" increased from universal (i.e. community samples; r = -0.104), selected (i.e. families with higher levels of risk factors for offspring DBDs; r = -0.134), indicated (i.e. families with emerging offspring DBDs; r = -0.265) and finally treatment prevention programmes (i.e. families (self-) referred to outpatient clinics; r = -0.326; Leijten et al., 2019). It may be that the smaller magnitude of effects in the current study reflects the characteristics of the samples included, which were predominantly community-based samples that have less severe levels of DBD symptoms. Therefore, the results may be more analogous to those found in RCTs of universal PMT programmes.

Finally, positive parenting had a small and non-significant effect on offspring DBD symptoms (pooled r = -0.06, *95% CI* = -0.15, 0.03). This result contrasts with RCTs identifying positive parenting practices as the key components of interventions (Leijten et al., 2019, 2022). Although our analyses for positive parenting were less powered than negative parenting (positive parenting SE [standard error] = 0.0459, negative parenting SE = 0.0194), the pooled estimate was derived from 16 studies, including 20,603 individuals. Furthermore, while non-significant, the results were in the hypothesised direction, with more positive parenting practices decreasing the risk of DBD symptoms. Emphasising positive parenting practices may be key in targeted prevention programmes (i.e., families with emerging or current DBD symptoms), whereas reducing negative parenting practices may be more important in universal prevention programmes where offspring exhibit fewer DBD symptoms.

Although the current study did not directly examine parenting interventions, the findings may prove useful for current interventions, especially universal prevention

efforts. Fathers are underrepresented in parenting interventions; our results indicate that the effect of parenting practices is similar for mothers and fathers. Therefore, paternal involvement in interventions may have a beneficial effect in preventing offspring DBD symptoms (Lundahl et al., 2008; Panter-Brick et al., 2014).

Our results suggest that for non-clinical samples, it is negative parenting practices, as opposed to positive parenting practices, that have a causal effect on DBD symptoms. Therefore, tailoring universal preventions to reduce negative parenting practices may be more beneficial.

The current study indicates a small causal effect of negative parenting practices, suggesting that there are many other causal risk factors for the development of DBDs. Other potential risk factors include peer deviance, parental psychopathology, and social disadvantage (Jaffee, Strait, et al., 2012). Future research should quantitatively synthesise the evidence for these factors via meta-analyses of studies using causal inference methods to triangulate evidence.

Finally, I estimate that even a small causal effect has the potential to substantially impact child behavioural development, with a 0.4SD reduction of negative parenting practices potentially leading to 3,614,337 school-aged children no longer displaying clinical levels of DBD symptoms. Due to the long-term adverse consequences of DBDs, preventing even a small fraction of the population from developing these symptoms is expected to have large and positive downstream consequences (Burt et al., 2018; Funder & Ozer, 2019).

4.4.1 Strengths and limitations

The current study is the first meta-analysis to consider the evidence for the relationship between parenting practices and offspring DBD symptoms from stringent causal inference methods. I used a broad definition of causal inference methods, which, on the one hand, meant that I captured a wide range of research articles to triangulate across different approaches but, on the other hand, may result in studies with different target populations, increasing between-study heterogeneity. I conducted sensitivity checks, including leave-one-out and subgroup analyses, to assess potential sources of heterogeneity and adjust the analyses accordingly. Furthermore, both parenting practices and DBD symptoms were mainly assessed via questionnaires, which can imprecisely capture the phenotypes they aim to measure and are prone to recall bias (see Chapter 1, Section 1.3.1.2.2). To strengthen the accuracy of the measures, I only included studies that used explicitly defined, e.g., well-validated, measures for the exposure and outcome. In addition, the findings from studies that used questionnaires were similar to those that used observational measures and semi-structured interviews, such as the Five-Minute Speech Sample (Gottschalk & Gleser, 1979).

I could not examine whether the findings were moderated by participant ancestry, as most studies did not report adequate information for subgroup analyses. In the studies that provided information on participant ancestry, most of the participants were White. Future research must find ways to improve diversity in research participation, and studies need to provide information on the ancestry of their samples for evidence and subsequent evidence-based practice to benefit all groups in society (Wellcome, 2021).

Finally, as mentioned above, although causal inference methods can more effectively control for potential confounders, they are not immune to bias. Most studies included

in this meta-analysis were family-based designs, such as discordant twin and sibling studies, which account for genetic confounding but do not always control for non-shared environmental confounding, such as differential experiences outside the home, unless adjusted for in the analyses. Although I cannot be certain that the included studies were not affected by unmeasured confounding, through comparing and combining methods which have different sources of bias, I can be more confident in our findings than if I considered one method alone (Goetghebeur et al., 2020; Lawlor et al., 2017; Munafò & Smith, 2018).

4.4.2 Conclusions

In this Chapter, I present results from my meta-analysis of studies using causal inference methods that suggest that negative parenting has a small harmful effect on DBDs. Interventions that target negative parenting practices could prevent approximately 3,614,337 cases of DBDs worldwide, which, given the high costs of disruptive behaviours, could have substantial benefits to individuals and the wider society. Future research using causal inference methods will be valuable in identifying other modifiable causes of DBDs, which, along with reducing negative parenting practices, could be incorporated into interventions for DBDs.

Key points

1. This current meta-analytic review suggests that negative parenting practices have a small, harmful effect on offspring DBDs.

2. Shared method variance should be considered when designing studies and assessing the findings of studies that used the same reporters for the exposure and the outcome.

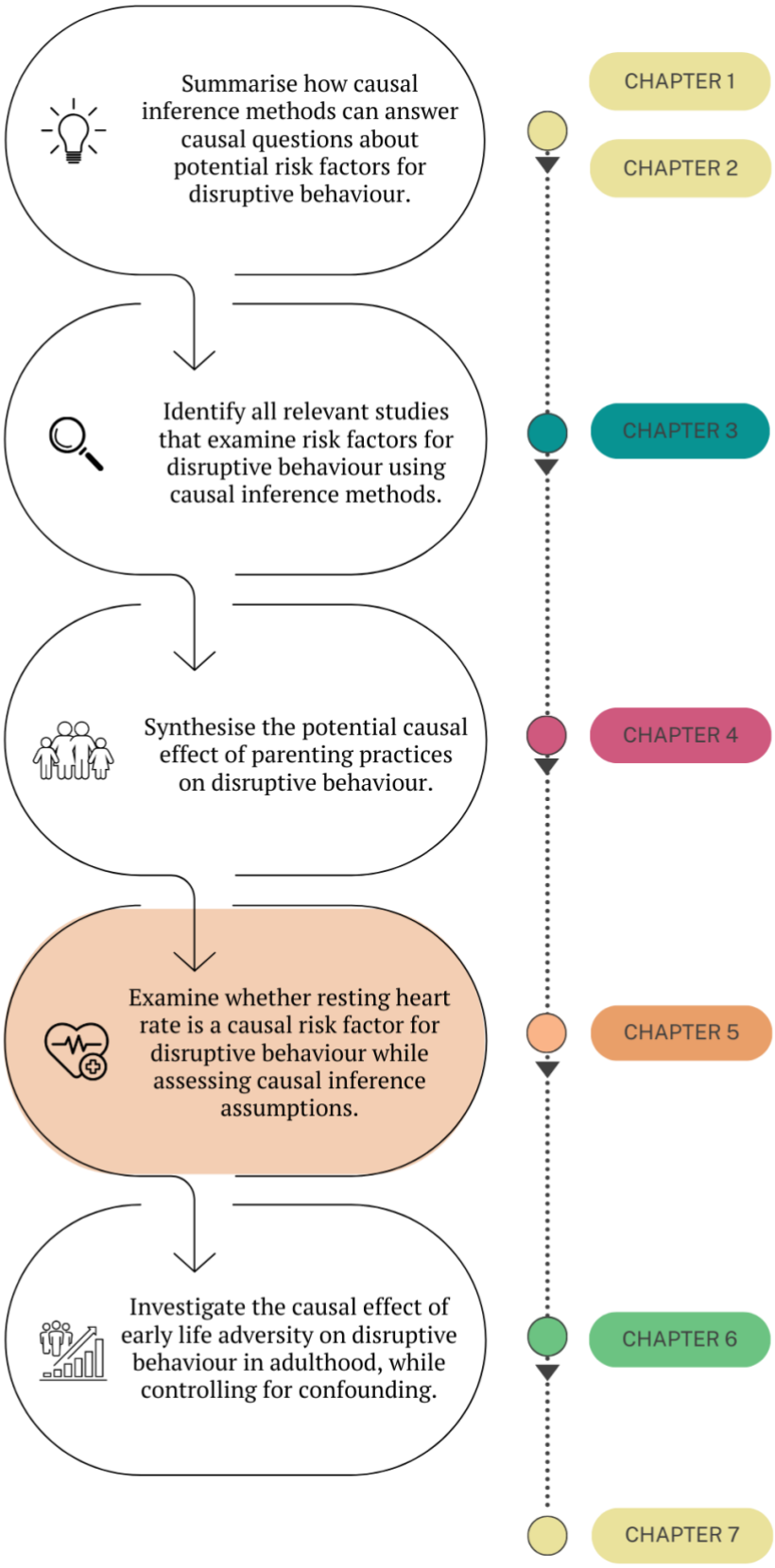
3. Results vary by type of causal inference method used as different methods have different estimands and control for different biases.

4. By evaluating the results of studies that use different causal inference methods, researchers can identify the potential confounders of the relationship between parenting and DBDs, including genetic and environmental confounders.

5. The population attributable impact may be useful for interpreting the pooled causal effect from meta-analyses of risk factors for DBDs.

6. In the current analyses, I estimated that prevention programmes which effectively target negative parenting practices could reduce the prevalence of DBDs by 0.11% worldwide, preventing over 3.5 million school-aged children from exhibiting clinical symptoms of disruptive behaviour.

THESIS STRUCTURE



5 RESTING HEART RATE AND ANTISOCIAL BEHAVIOUR: A MENDELIAN RANDOMISATION STUDY

Chapter overview

In Chapter 4, I conducted the first systematic review and meta-analysis on evidence for parenting practices and disruptive behaviour from studies using causal inference methods. I chose to focus on parenting practices as I found in my review of all risk factors (Chapter 3) that previous studies had used a wide variety of methods to investigate this risk factor. However, I did not identify any studies that used causal inference methods to investigate psychophysiological risk factors for DBDs. In this Chapter, I will explore how Mendelian randomisation analyses and linkage disequilibrium score regression analyses can be applied when studying biological risk factors for DBDs. I will focus on one of the most common psychophysiological risk factors associated with DBDs: resting heart rate. I will assess the causal inference assumptions associated with instrumental variable analyses while examining whether resting heart rate is causally related to disruptive behaviour.

Publication status: This Chapter is based on an article published in *Scientific Reports*: Karwatowska, L., Frach, L., Schoeler, T., Tielbeek, J. J., Murray, J., de Geus, E., Viding, E., & Pingault, J.-B. (2023). Resting heart rate and antisocial behaviour: A Mendelian randomisation study. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-37123-y

5.1 Background

Antisocial behaviour (ASB) is a common symptom of disruptive behaviour disorders (DBDs) and includes behaviours such as aggression, rule-breaking and acts of violence. Individuals who display high levels of ASB are at risk of lifelong adverse outcomes, such as poor mental health, substance misuse, criminal behaviour, and unemployment (Bevilacqua et al., 2018; Colman et al., 2009; Piquero et al., 2007). Considering these long-term and pervasive adverse outcomes, it is important to understand the aetiology of ASB to inform early identification and evidence-based intervention efforts.

Numerous reviews exist on putative risk factors for ASB, which include environmental and neurobiological factors (Derzon, 2010; Hawkins et al., 2000; Jaffee, Strait, et al., 2012; Murray & Farrington, 2010). Physiological markers, such as those indexing autonomic nervous system (ANS) activity, are particularly important in elucidating potential mechanisms underlying the development of ASB (Fanti, 2018; Matthys et al., 2013). Of these, resting heart rate (RHR), defined as the number of heart beats per minute while at rest, is the most well-studied. Observational studies frequently report a strong inverse relationship between RHR and ASB, where individuals with lower RHR display higher levels of various types of ASB, including child conduct problems (CP), juvenile delinquency, and adult violence (Baker et al., 2009; Bergstrøm & Farrington, 2018; J. R. Jennings et al., 2017; Raine et al., 1997; Schoorl et al., 2016; Wadsworth, 1976). This evidence has been synthesised in several meta-analyses, all reporting a robust association between the two traits (de Looff et al., 2022; Lorber, 2004; Ortiz & Raine, 2004; Portnoy & Farrington, 2015), with one stating that RHR is a “possible causal risk factor for antisocial behavior” (Portnoy et al., 2015).

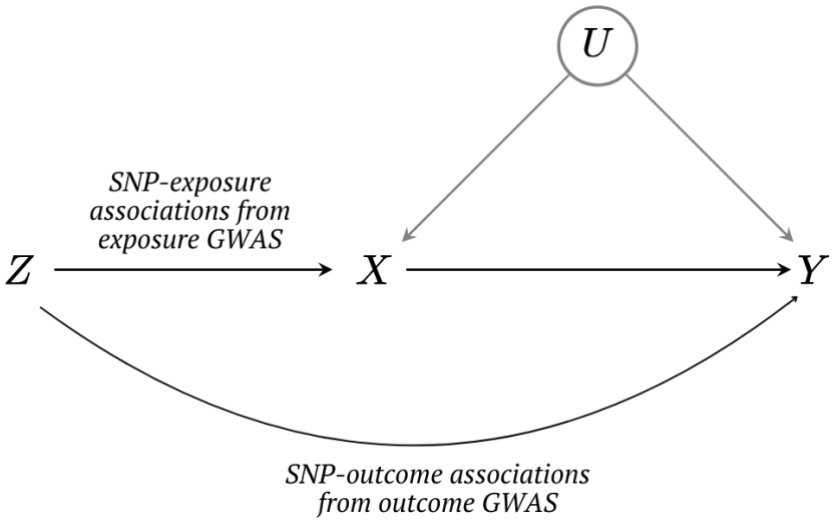
Various potential mechanisms have been proposed to explain the relationship between RHR and ASB. The two main theories are the fearlessness (Raine, 1993) and sensation-seeking hypotheses (Raine, 2002). According to the fearlessness hypothesis, an individual with a low RHR has a higher threshold for experiencing fear than individuals with higher RHRs, partly due to attenuated ANS responses to aversive stimuli. The typical links between poor behavioural choices (e.g., aggression) and aversive stimuli (e.g., perceived punishment cues) are either not established or are insufficiently established in individuals with lower RHR. As such, an individual with lower RHR would have inappropriately low expectations of negative outcomes and be prone to repeat poor decision-making. In support of this hypothesis, many behavioural experiments report that participants who show deficient fear conditioning and reduced anticipatory fear reactivity have lower RHRs and higher levels of ASB (Gao et al., 2010).

The sensation-seeking hypothesis states that individuals with lower RHR have low basal ANS activity and are chronically hypo-aroused. Hypo-arousal is an unpleasant physiological state, and therefore, individuals with lower RHR seek to increase their arousal to a normal level by engaging in ASB. Sensation-seeking is associated with both RHR and ASB, with some evidence suggesting that sensation-seeking is a mediator of these two factors (Hammerton et al., 2018; Portnoy et al., 2014; Sijtsema et al., 2010).

Although the link between RHR and ASB has been well studied, questions remain about whether these two phenotypes are causally related. This is partly due to limitations in the existing literature that prevent the drawing of causal conclusions. A closer look at the studies included in the four meta-analyses (de Looff et al., 2022; Ortiz & Raine, 2004; Portnoy & Farrington, 2015) shows that most studies have used small and/or selective samples, which can produce unreliable and un-generalisable

estimates. In recent years, authors have attempted to include larger, unselected samples, which have been followed up over time. The findings from these studies are more variable than those from earlier studies. Indeed, more recent studies have either confirmed earlier findings of a strong negative relationship between RHR and ASB (Armstrong et al., 2019; Hammerton et al., 2018; Latvala et al., 2015; Murray, Hallal, et al., 2016) or found no relationship between the two factors (Fanti, 2018; Fanti et al., 2019; Kavish et al., 2019; Oldenhof et al., 2019; Prätzlich et al., 2019). Furthermore, most of the findings of a strong relationship between RHR and ASB are from associational studies. Indeed, there is a lack of research adopting causal inference approaches to help overcome the inherent biases of these studies. For example, only two studies have used genetically informed family-based methods, and both found that the relationship between heart rate and ASB is entirely explained by genetic effects, i.e., genetic confounding (Baker et al., 2009; Kendler et al., 2021). Interestingly, these studies found evidence of genetic covariation between RHR and ASB, whereby children with a genetic liability for lower RHR also have a genetic liability for ASB.

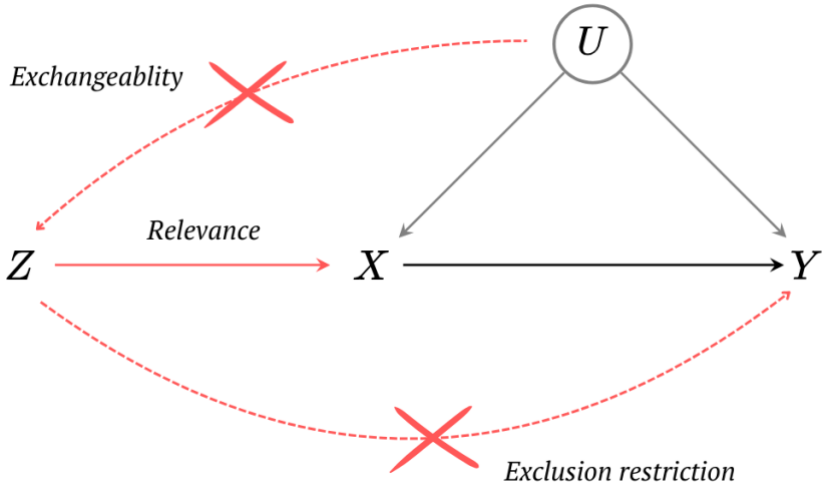
Genetically informed methods can exploit the heritability of both RHR (de Geus et al., 2007; Eppinga et al., 2016) and ASB (Lewis & Plomin, 2015; Polderman et al., 2015; Salvatore & Dick, 2018). The two aforementioned genetically informed studies used methods which rely on knowledge of genetic relatedness between family members. Other genetically informed methods can be used to triangulate these findings by relying on different types of data and assumptions (Lawlor et al., 2017; Munafò & Smith, 2018). A useful genetically informed causal inference method is Mendelian randomisation (MR). MR is an instrumental variable approach that uses genetic variants (i.e., single nucleotide polymorphisms; SNPs) associated with an exposure of interest (e.g., SNPs associated with RHR) as instrumental variables (IVs) to assess the effect of an exposure of interest (e.g., RHR) on an outcome (e.g., ASB; see Figure **5.1**).



Note. Z = instrumental variable (single nucleotide polymorphisms, X = exposure, Y = outcome, U = unmeasured confounders.

Figure 5.1 *An illustration of two-sample Mendelian randomisation analyses, which uses summary-level data from two separate genome-wide association studies: one for the exposure (to estimate the SNP-exposure associations) and one for the outcome (to estimate the SNP-outcome associations).*

MR provides causal effect estimates under the classic IV assumptions (Chapter 1, Section 1.4.2.1.2), illustrated in Figure **5.2**. The genetic variants indexing the exposure must be *(a)* associated with the exposure (*relevance*), *(b)* independent of confounders of the exposure-outcome relationship (*exchangeability*), and *(c)* only associated with the outcome through the exposure (*exclusion restriction*; Burgess et al., 2017; Davey Smith & Hemani, 2014).



Note. Z = instrumental variable (single nucleotide polymorphisms, X = exposure, Y = outcome, U = unmeasured confounders.

Figure 5.2 *An illustration of the main Mendelian randomisation assumptions that the instrument is associated with the exposure (relevance), is not associated with (un)measured confounders (exchangeability) and does not influence the outcome other than through the exposure (exclusion restriction).*

MR analyses use SNPs as IVs in a similar way as RCTs use the randomisation procedure of interventions as an IV of intervention uptake. This is possible due to Mendel's second law, also known as the law of independent assortment, which states that genetic variants are randomly assigned at conception (Davey Smith & Ebrahim, 2003). However, it is important to note that although most genetic variants are randomly assigned, some genetic variants are in linkage disequilibrium (LD: i.e. they are co-inherited, which induces non-random correlations between SNPs). LD and pleiotropy (i.e. when a single SNP affects more than one trait) can bias MR analyses. Therefore, efforts must be made to ensure that results are not affected by the presence of one or both of these conditions (Lawlor et al., 2008). As MR analyses depend on a different set of assumptions and use different types of data to other causal inference methods, they provide a quick and efficient way to triangulate evidence on potentially modifiable risk factors (Sanderson, Glymour, et al., 2022).

It is not known whether the association between RHR and ASB reflects causality. The current study will be the first to interrogate the potential causal effect of RHR on ASB using two-sample MR analyses. I will exploit powerful genetic data from two large, independent genome-wide association studies (GWAS) on RHR (Zhu et al., 2019) and ASB (Tielbeek et al., 2022) to provide a new type of evidence for triangulation with previous observational studies on this topic. Most evidence suggests an association between lower RHR and higher ASB, although research using more rigorous approaches calls into question the strength of this association. The current study is exploratory and, therefore, had no prior hypotheses. Instead, this study aimed to investigate whether RHR has a causal effect on ASB.

5.2 Methods

5.2.1 Study design

To identify potential causal effects of resting heart rate (RHR) on antisocial behaviour (ASB), I conducted two-sample Mendelian randomisation (MR) analyses using multiple single nucleotide polymorphisms (SNPs) as instrumental variables (IVs) for RHR (Burgess et al., 2013, 2015, 2020; Davey Smith & Ebrahim, 2003; Davey Smith & Hemani, 2014; Pierce & Burgess, 2013). Univariable two-sample MR integrates summary level genetic data from two GWAS, one GWAS estimating the association between multiple SNPs and the exposure (i.e., SNPs-RHR), and the other, independent, GWAS estimating the associations between the same SNPs and the outcome (i.e., SNPs-ASB). I used the inverse-variance weighted (IVW) estimate estimator, which can be interpreted as the weighted average of the effect estimates across all SNPs (i.e., each SNP provides one estimate of the causal effect of interest; Burgess et al., 2013). The IVW estimator is the estimate from a weighted regression of the SNP-outcome effects on SNP-exposure effects where the intercept is constrained to zero, and the weighting is based on the inverse of the variance of the

SNP–outcome association, thereby reflecting the precision of each instrument (i.e. each SNP).

I ran several sensitivity analyses to test for potential violations of the MR assumptions (Burgess et al., 2017; Davey Smith & Hemani, 2014). These included other MR methods such as MR Egger (Bowden et al., 2015), weighted median analysis (Bowden, Davey Smith, et al., 2016), MR Robust Adjusted Profile Score (RAPS; Zhao et al., 2020), MR Pleiotropy RESidual Sum and Outlier (PRESSO; Verbanck et al., 2018) and contamination mixture methods (Burgess et al., 2020). I further checked for heterogeneity and horizontal pleiotropy using the MR Egger intercept and used the Steiger approach to rule out reverse causation (Hemani et al., 2017). A dictionary that provides a comprehensive and accessible overview of MR theory, methodology and interpretation is available online (Lawlor et al., 2019).

The autonomic nervous system (ANS) controls both RHR and resting heart rate variability (HRV). To assess whether another associated measure of ANS activity drives the association between RHR and ASB, I conducted further analyses, including a multivariable MR analysis using data from a GWAS of HRV (Nolte et al., 2017). Multivariable MR is an extension of univariable MR, which, instead of using SNPs for one exposure, includes SNPs for multiple exposures, thereby assessing the causal effect of related exposures on the outcome simultaneously (Burgess et al., 2013). By adding SNPs for HRV, I better capture the individual differences in ANS functioning underlying low RHR, in particular, high levels of cardiac vagal control. As such, I could assess the effects of RHR on ASB independent of cardiac vagal control, for instance, due to cardiac sympathetic control, and therefore elucidate mechanisms of ANS activity. As a positive control, I conducted MR between RHR and HRV as an alternative outcome, assuming a negative causal effect. I also conducted linkage disequilibrium (LD) score regression (LDSC) to estimate genetic correlations between

the heart rate measures and ASB from the GWAS summary statistics (LDSC; Bulik-Sullivan et al., 2015, 2015). LDSC assesses whether SNPs are in LD and, therefore, whether confounding is leading to inflated false positives.

I followed the Strengthening the Reporting of Observational Studies in Epidemiology– Mendelian Randomization (STROBE-MR) guidelines (Skrivankova et al., 2021; see Appendix D, Table 1).

5.2.2 Data sources

The main analyses used data from two summary statistics files, one for RHR (*n* = 458,835) and one for ASB (*n* = 85,359), with no sample overlap in the GWAS. In the secondary analyses, which included summary statistics for HRV, there was limited overlap between the HRV (*n* = 53,174) and ASB GWAS, with a potential overlap of 1,300 individuals. Further information on the cohorts and measures used in all the original GWAS is available in Table **5.1** and Table **5.2**. Ethical approval for each GWAS was obtained by the authors of the original studies (Nolte et al., 2017; Tielbeek et al., 2022; Zhu et al., 2019).

5.2.3 Measures

5.2.3.1 Exposure

The summary statistics for RHR were obtained from the largest and most recent GWAS on RHR (Zhu et al., 2019). The GWAS included 458,835 individuals from the UK Biobank (Allen et al., 2014) and in addition to age and sex, which was also controlled for in the outcome GWAS (below), the authors also controlled for smoking. As smoking is associated with ASB and was not controlled for in the ASB GWAS, the authors agreed to rerun the analysis without controlling for smoking. Heart rate was measured during sitting at rest for 2-3 minutes. The reading was taken during blood

pressure measurement and/or using the pulse waveform obtained from the finger with an infrared sensor during arterial stiffness measurement. An average was taken where multiple RHR measurements were available for one individual. The exposure is expressed in beats per minute.

5.2.3.2 Outcome

Summary statistics for ASB were obtained from the Broad Antisocial Behavior Consortium (BroadABC) GWAS, which includes 85,359 individuals from 28 discovery and five independent replication samples (Tielbeek et al., 2022). The ASB measures from these samples covered a broad range of behaviours, including conduct disorder, aggression, and delinquency (see Table **5.1** for further information on cohort demographics and the types of ASB assessed). The measures were used to derive a single quantitative measure for ASB. The outcome was a standardised score on a continuous trait scale. Like the exposure GWAS, age and sex were adjusted for in the meta-analyses.

Table 5.1 *Information on cohorts included in the genome-wide association studies used to access the genetic variants associated with resting heart rate and antisocial behaviour in the main analyses.*

Cohort	Study design	Measures	Sample size
Resting heart rate (*n* = 458,835)			
UK Biobank	Longitudinal	Automated reading during blood pressure measurement; pulse waveform obtained from the finger with an infrared sensor during arterial stiffness measurement.	458,969
Antisocial behaviour (*n* = 85,359)			

Cohort	Study design	Measures	Sample size
The National Longitudinal Study of Adolescent to Adult Health (AddHealth)	Longitudinal	Aggression – questionnaire on violent and non-violent activities	5,874
Avon Longitudinal Study of Parents and Children (ALSPAC)	Longitudinal birth cohort	Antisocial behaviour – Edinburgh Study of Youth Transitions and Crime	2,942
Brain Imaging Genetics (BIG)	Population-based	Aggression – Reactive–Proactive Aggression Questionnaire (RPQ)	862
Collaborative Study on the Genetics of Alcoholism (COGA)	Clinically based family study	Antisocial personality disorder – Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)	7,274
CoLaus|PsyCoLaus	Population-based	Antisocial personality disorder – Semi-structured Diagnostic Interview for Genetic Studies (DIGS)	4,071
Finnish Twin Cohort (FinnTwin)	Population-based	Antisocial personality disorder – Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)	2,554
The Genetics of Sexuality and Aggression (GSA)	Population-based	Aggression – Buss & Perry Aggression Questionnaire	2,329
Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH)	Case-control	Antisocial personality disorder – ICD-10	24,819
Minnesota Center for Twin and Family Research (MCTFR)	Family study	Antisocial personality disorder – Structured Clinical Interview for DSM-IIIR Disorders (SCID)	5,943
Phenomics and Genomics Sample (PAGES)	Population-based	Aggression – Buss-Durkee Hostility	2,480

Cohort	Study design	Measures	Sample size
		Inventory (Assault and Aggression subscales)	
Psychiatric Genomics Consortium (Cardiff sample, CHOP cohort, IMAGE-I & IMAGE-II samples, Barcelona sample, Yale-Penn cohort)	Case-control	Various	10,288
QIMR Berghofer Medical Research Institute (QIMR)	Population-based family study	Aggression – Buss-Durkee Hostility Inventory (Assault and Aggression subscales) Antisocial personality disorder – questionnaire based on DSM-IV criteria Arrest – count of number of arrests since 18th birthday	10,363
Spit for Science (S4S)	Population-based	Antisocial behaviour – Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)	2,187
Twins Early Development Study (TEDS)	Population-based	Antisocial personality disorder – adapted from Edinburgh Study of Youth Transitions and Crime	3,694
Tracking Adolescents' Individual Lives Survey (TRAILS)†	Population-based	Antisocial behaviour – Antisocial Behavior Questionnaire	1,360

Note. † Shared between the heart rate variability and antisocial behaviour GWAS.

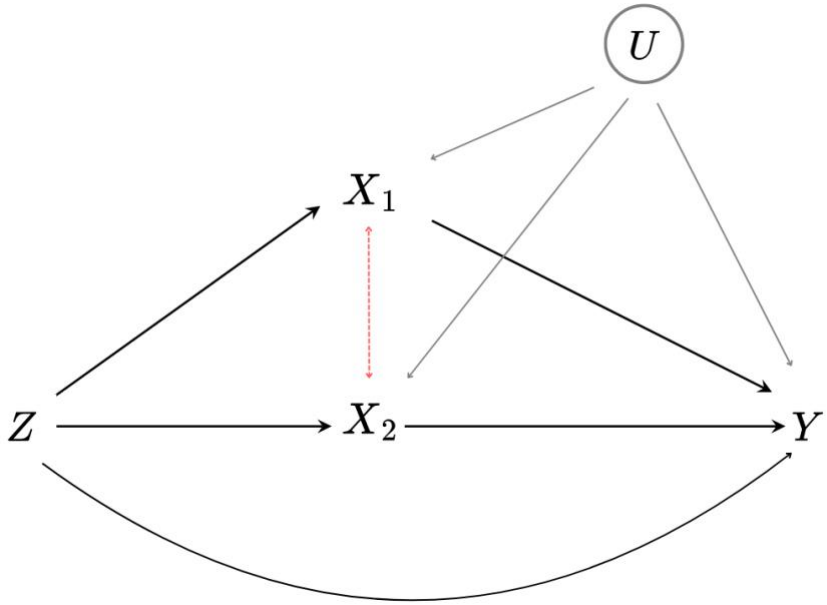
5.2.4 SNP selection

Prior to the main analyses, I conducted quality control procedures on the GWAS summary statistics. For the exposure GWAS to align with the outcome GWAS, I retained autosomal SNPs (i.e., SNPs located on non-sex chromosomes) with a minor allele frequency > 0.01 and an imputation information (INFO) score > 0.6. To fulfil the *relevance* assumption (i.e. the instruments are associated with the exposure), I used SNPs associated with the exposure above the genome-wide significance level of $p < 5 \times 10^{-8}$. The selected SNPs were then clumped for independence using default parameters from the R package *"TwoSampleMR"* (Hemani, Zheng, et al., 2018), excluding any SNPs with a pairwise $r^2 > 0.001$ within a 10000 kilobase window. The genome-wide significant, independent SNPs for RHR were then harmonised with SNPs from the ASB GWAS, retaining variants present in both GWAS.

5.2.5 Sensitivity analyses

5.2.5.1 Multivariable MR with heart rate variability

Resting heart rate variability (HRV) captures the vagal effects on the sinoatrial node co-determining RHR and has also been considered a potential cause of ASB (Beauchaine et al., 2019; Beauchaine & Thayer, 2015). Therefore, to rule out alternative explanations and potential causes from other related measures of ANS activity, I ran multivariable MR of RHR and HRV on ASB. In multivariable MR, SNPs associated with at least one exposure (i.e. RHR or HRV) are included in the MR analyses (see Figure **5.3**). Thereby simultaneously estimating each exposure's "direct" causal effect, conditional on the other exposure(s).



Note. Z = instrumental variable (single nucleotide polymorphisms, X_1 = exposure, X_2 = second exposure, Y = outcome, U = unmeasured confounders.

Figure 5.3 *An illustration of multivariable Mendelian randomisation analyses with two exposures. The single-nucleotide polymorphisms are associated with at least one of the exposures. The line between X_1 and X_2 is left bidirectional and dashed, as no assumptions are made about this relationship.*

I obtained summary statistics for resting HRV from the Genetic Variance in Heart Rate Variability (VgHRV) Consortium GWAS (Nolte et al., 2017) of 53,174 participants. HRV was measured using three traits: the standard deviation of the normal-to-normal inter-beat intervals (SDNN); the root mean square of the successive differences of inter-beat intervals (RMSSD); the peak-valley respiratory sinus arrhythmia or high-frequency power (pvRSA/HF). All three traits were measured during resting, basal recordings either via electrocardiograms (10-s, 20-s, up to 90 min of sitting or 2 – 12-hour daytime), 24hr Holter monitor, finger photoplethysmography or Portapres ambulatory heart rate recordings (Nolte et al., 2017). See Table **5.2** for more information. The same clumping and harmonisation parameters used for the main analyses, i.e. the univariable MR of RHR, were also used for the additional analyses, i.e. the multivariable MR including RHR and HRV.

Table 5.2 *Information on cohorts included in the genome-wide association study used to access the genetic variants associated with heart rate variability in the supplementary analyses.*

Cohort	Study design	Measures	Sample size
Heart rate variability (*n* = 53,174)			
Atherosclerosis Risk in Communities Study (ARIC)	Population-based	3-lead ECG, 2 minutes, supine (RMSSD; SDNN; HF)	8,262
Cardiovascular Health Study (CHS)	Population-based	24hr Holter monitor (RMSSD; SDNN; HF)	8,262
Framingham Heart Study (FHS)	Population-based	2hr ambulatory ECG (RMSSD; SDNN; HF)	1,944
FINnish Genetic STUdy of aRrhythmic Events (FINGESTURE)	Prospective case-control study	24hr Holter monitor (SDNN; HF)	494
FLEMish study on Environment, Genes and Health Outcomes – European Project on Genes in Hypertension (FLEMENGHO-EPOGH)	Population-based	12-lead ECG & nasal thermistor for RSA: PSA to estimate HF ranges; ECG recording for 15 min; supine (pvRSA)	196
Generation R Study (GenR)	Population-based	3-pole ECG & breathing pattern using a piezoelectric transducer; 100-180 seconds; sitting (HF)	392
Groningen Twin Registry (GTR)	Twin study	Type II 3-lead ECGs & respiration with a flexible band around upper thorax; 5 minutes; sitting (RMSSD; SDNN; HF)	134
Kooperative gesundheitsforschung in der Region (KORA S4)	Population-based	2-lead ECG; 5 minutes; supine (RMSSD; SDNN; HF)	1,617
Multi-Ethnic Study of Atherosclerosis (MESA)	Population-based	12-lead ECG; average from 3 sequential 10-second ECGs; supine; resting (RMSSD; SDNN)	2,401
Marine Resiliency Study (MRS)	Population-based	Finger photoplethysmography; 5 minutes; sitting (RMSSD; SDNN; HF)	1,383

Cohort	Study design	Measures	Sample size
Netherlands Study of Depression and Anxiety (NESDA)	Case-control study	Type II, 3-lead ECG & breathing recorded from thorax impedance; ~90 minutes; sitting (RMSSD; SDNN; pvRSA)	1,740
Netherlands Twin Register (NTR)	Twin-family study	Type II, 3-lead ECG & breathing recorded from respirate; 8 minutes; sitting (RMSSD; SDNN; pvRSA)	439
Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	Population-based	6-precordial-lead ECG & breathing recorded using custom- made impedance device; 5-minutes; supine; controlled breathing (12 breaths/min; SDNN; pvRSA)	766
Prevention of Renal and Vascular End- stage Disease (PREVEND)	Population-based	Beat-to-beat blood pressure pulse wave recording on middle finger (Portapres); 15 minutes; supine (RMSSD; SDNN; HF)	2,793
Rotterdam Study (RS1+2)	Population-based	12-lead ECG; 10 seconds; resting (RMSSD; SDNN)	972 + 985
Tracking Adolescents Individual Lives Survey – CliniCal cohort (TRAILS-CC)	High-risk adolescent cohort	Type II 3-lead ECG; 4 minutes (T1); supine (RMSSD; SDNN; HF)	307
Tracking Adolescents Individual Lives Survey – POPulation cohort (TRAILS-Pop) †	Population-based	type II 3-lead ECG; 4 minutes (T1), 5 minutes (T3); supine (RMSSD; SDNN; HF)	1,222
Uppsala Longitudinal Study of Adult Men (ULSAM)	Population-based	6-precordial-lead ECG & breathing recorded using custom- made impedance device from a 24hr recording during normal activity (SDNN; pvRSA)	67
Cardiovascular Risk in Young Finns Study (YFS)	Population-based	2-lead ECG; 3 minutes; supine (RMSSD; SDNN; HF)	1,827

Note. †Cohort shared between the heart rate variability and antisocial behaviour GWAS. Abbreviations: RMSSD = the root mean square of the successive differences of inter-beat intervals; SDNN = the standard deviation of the normal-to-normal inter-beat intervals; pvRSA/HF = the peak-valley respiratory sinus arrhythmia or high-frequency power.

5.2.5.2 LD Score Regression

To estimate the genetic correlation between all heart rate measures and ASB, I conducted LDSC (Bulik-Sullivan et al., 2015, 2015). LDSC is run in the command line using Python code from the open-source "*munge_sumstats.py*" script available on GitHub (repository: bulik/ldsc; Bulik-Sullivan et al., 2015, 2015). LDSC was implemented using the default parameters, including filtering the summary statistics to remove palindromic SNPs (i.e. SNPs whose alleles are the same on both strands and are therefore difficult to harmonise), removing SNPs with INFO scores (a measure of how well a SNP has been imputed) < 0.9 and removing SNPs which are available for less than 67% of individuals (i.e. rare genetic variants).

5.2.6 Statistical analyses

Apart from the LDSC analyses, which were conducted using a publicly available command line tool, I performed all analyses in *R (*R Core Team, 2019) using the *TwoSampleMR* (Hemani, Zheng, et al., 2018) and *MendelianRandomization* (Yavorska & Burgess, 2017) packages.

5.3 Results

5.3.1 Univariable MR analyses between resting heart rate and antisocial behaviour

After harmonisation, 300 genetic variants were available in both exposure and outcome datasets. Eight palindromic SNPs were removed. For the remaining 292 variants, I investigated the direction of their effects using Steiger filtering, which ensures that the direction of effects between the exposure and outcome are not

misclassified by removing SNPs that explain more variation in the outcome than the exposure. Fourteen SNPs showed higher associations with the outcome than the exposure and were removed, leaving 278 variants. The MR analyses of RHR on ASB did not support a causal effect, with 95% confidence intervals including the null value (N_{SNPs} [number of SNPs]= 278; B_{IVW} [inverse variance weighted estimate]= –0.0004; *95% CI* = -0.004, 0.004; p = 0.841; Table **5.3**). The scatterplot (Figure **5.4**) shows that the IVW estimator is consistent with a null effect.

Table 5.3 *Results from the univariable Mendelian randomisation analyses on resting heart rate and antisocial behaviour.*

Method	**N_{SNPs}**	**B**	**SE**	***lCI***	***uCI***
IVW	278	–0.0004	0.002	-0.004	0.004
MR Egger	278	–0.0007	0.004	-0.008	0.007
Weighted median	278	–0.0039	0.003	-0.010	0.002
MR RAPS	278	–0.0009	0.002	-0.005	-0.003
MR PRESSO	278	–0.0004	0.002	-0.004	0.003
Contamination mixture	278	–0.0023	0.003	-0.008	0.004

Abbreviations: N_{SNPS} = number of single nucleotide polymorphisms, B = beta coefficient, SE = standard error, lCI = lower bound of 95% confidence interval, uCI = upper bound of 95% confidence interval, IVW = inverse variance weighted, MR = Mendelian randomisation, RAPS = Robust Adjusted Profile Score, PRESSO = Pleiotropy RESidual Sum and Outlier.

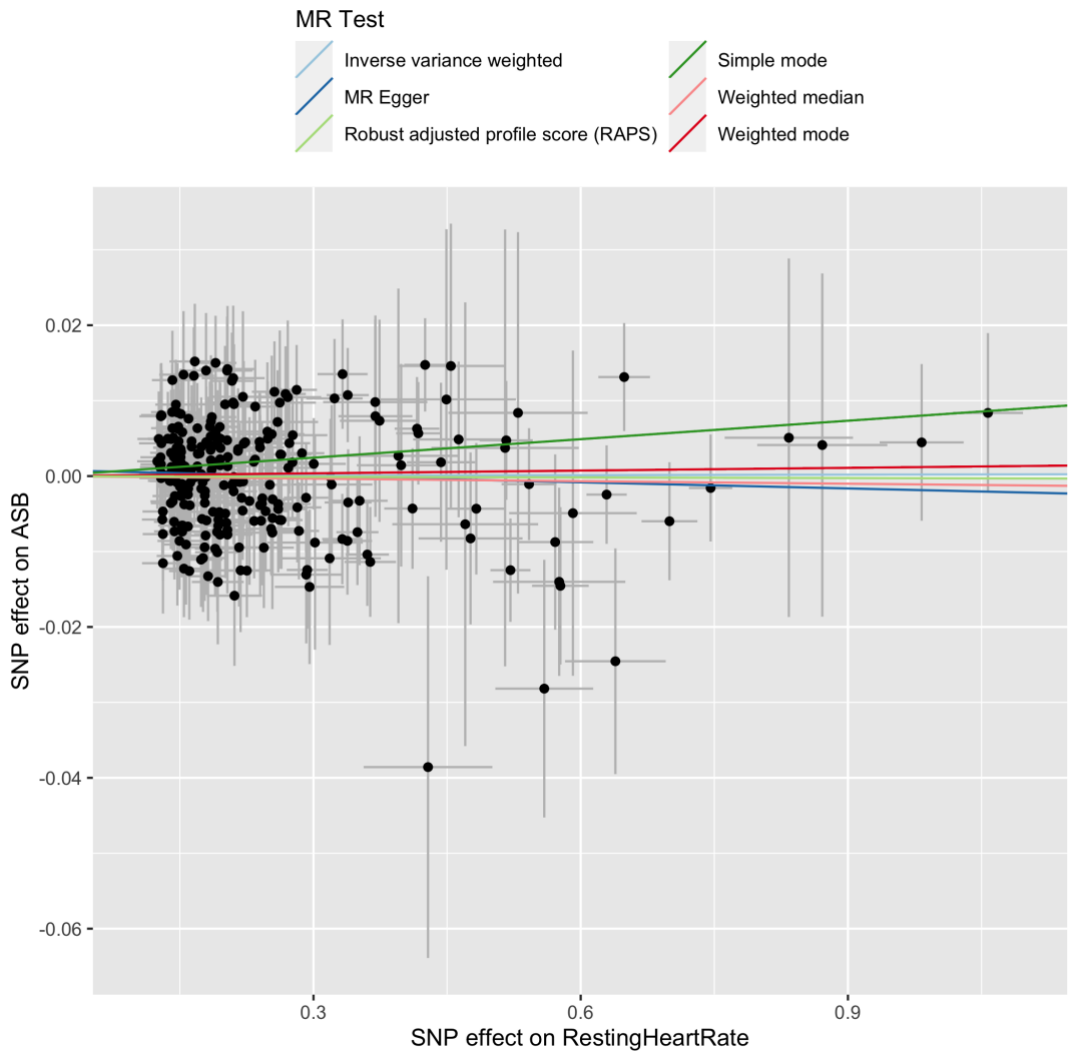


Figure 5.4 *A scatter plot of the association between the SNP effects on resting heart rate and SNP effects on antisocial behaviour.*

5.3.2 Sensitivity analyses

Besides the IVW estimator, I applied five other MR methods. These estimators are also summarised in Table **5.3**. Apart from the MR RAPS estimator, all other estimates included the null value. MR RAPS is often used to replicate the findings when other estimators indicate causal effects, as it accounts for weak instrument bias, pleiotropy, and extreme outliers. However, in the current analyses, no other estimator indicated causal effects, and, therefore, the MR RAPS estimator was not overinterpreted.

Regarding heterogeneity, the IVW and MR Egger Cochrane's *Q* statistics revealed significant heterogeneity in the estimates of the SNP-outcome associations (Q = 339.32, p = 0.024 and Q = 339.32, p = 0.027 respectively). This heterogeneity is visible in the forest plot of the effects of RHR on ASB (Figure **5.5**). In the absence of heterogeneity, the estimates would be relatively similar. However, if heterogeneity is present, then the estimates are variable.

No directional pleiotropic effects were observed from the MR Egger intercept (intercept = 0.0000185; p = 0.985) or MR PRESSO global test for pleiotropy (RSS_{obs} = 252.642, p = 0.872), which supports the exclusion restriction assumption.

The MR Steiger test revealed that associations between the genetic instruments and the exposure were 7.49 times higher than between the genetic instruments and the outcome (R^2_{EXP} = 0.05, R^2_{OUT} = 0.007). Therefore, there was no evidence of reverse causation.

The results did not change when conducting leave-one-out analyses using IVW (p_{min} = 0.605) and MR Egger (p_{min} = 0.598), indicating no influential outliers. The genetic instruments had a high average *F* statistic (F = 84, range = 27, 1185), suggesting that the assumption of no measurement error held. Evidence of this is further by the I^2 statistic of the SNP-exposure association (I^2_{GX}), which lies between 0 and 1 and indicates whether the no measurement error assumption has been violated. Scores closer to 1 indicate limited bias in the estimate. The I^2_{GX} statistic in the current analyses was 0.99, which further indicates that the results were not affected by weak instrument bias (Bowden, Del Greco M, et al., 2016).

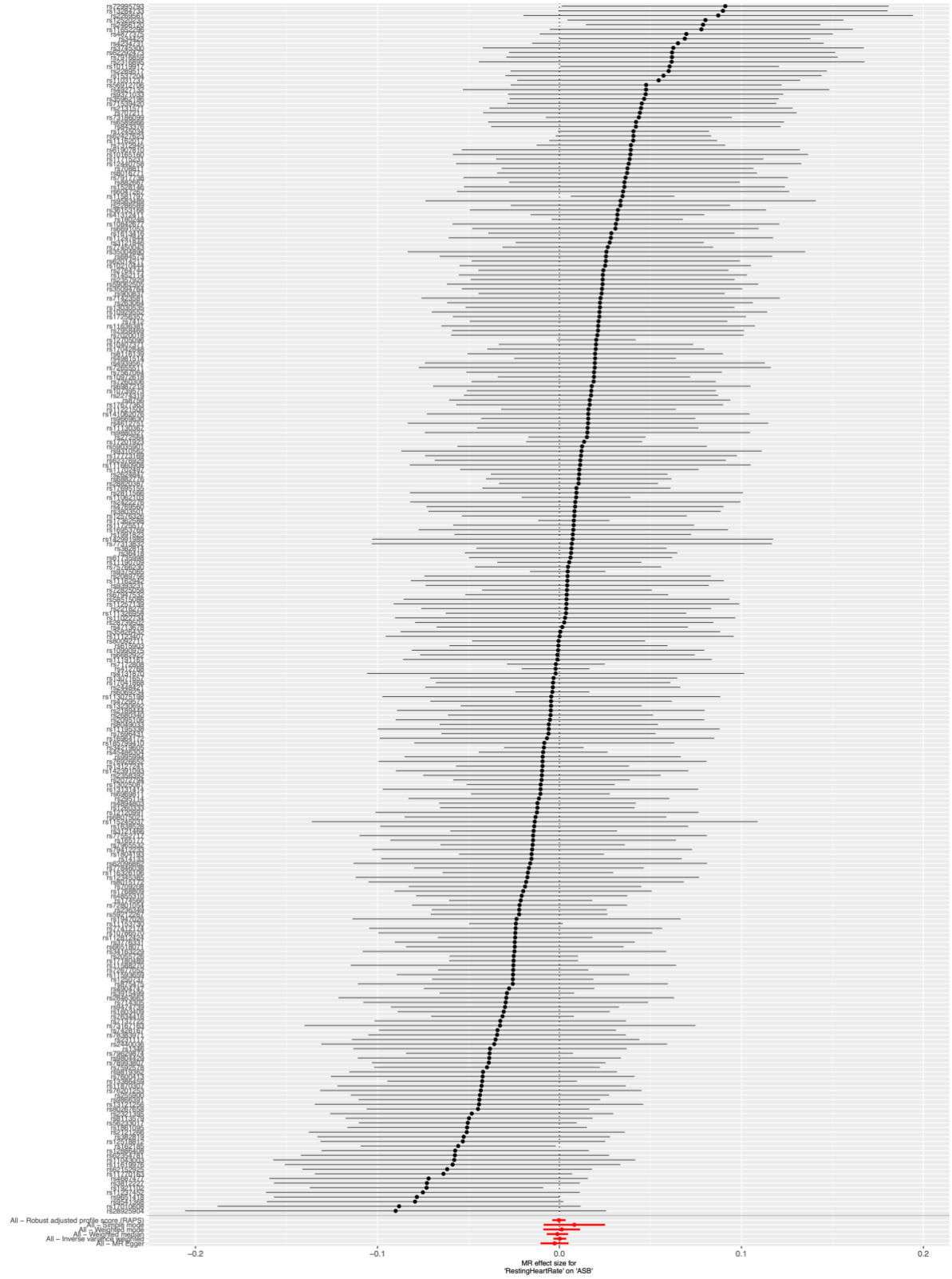


Figure 5.5 *A forest plot of the association between independent SNPs for resting heart rate and antisocial behaviour.*

5.3.3 Univariable MR analysis between heart rate variability and antisocial behaviour

To investigate alternative explanations for the association between RHR and ASB reported by previous research, I conducted additional analyses, including univariable MR with HRV, multivariable MR and LD score regression. The results from the univariable MR analyses suggested no causal effect of any of the three measures of HRV on ASB (Table **5.4**). These results were not affected by heterogeneity, as indicated by the IVW and MR Egger estimates. The MR Egger intercept and MR PRESSO estimate suggest no pleiotropic effects, and there was no evidence of weak instrument bias (*F* statistic range = 57 - 86; all I^2_{GX} = 0.99). MR Steiger filtering showed that the SNP effects on HRV were between 61 and 121 times higher than the SNP effects on ASB.

As the sample size of the underlying GWAS for HRV was relatively small, and only a few significant SNPs were detected, I also repeated the HRV analyses using a more liberal *p*-value threshold of $p < 5e\text{-}05$. The results were comparable with those from the initial analyses using the p-value threshold of $p < 5e\text{-}08$. Of note, although the MR Egger estimate for HRV (measured by RMSSD) on ASB indicated possible causal effects (N_{SNPs} = 52, *B* = 0.109, SE = 0.045, *p* = 0.019), the MR Egger intercept suggested the presence of directional pleiotropy (intercept = –0.005, SE = 0.002, *p* = 0.012). Therefore, I only report the estimates from the analyses using the more stringent p-value threshold of $p < 5e\text{-}08$ (Table **5.4**).

Table 5.4 *Results from the multiple univariable Mendelian randomisation analyses on heart rate variability and antisocial behaviour.*

Method	N_{SNPs}	*B*	*SE*	*p*
RMSSD				
IVW	5	0.039	0.072	0.586
MR Egger	5	0.179	0.171	0.375
Weighted median	5	0.090	0.079	0.251
MR RAPS	5	0.060	0.066	0.363
MR PRESSO	5	0.039	0.072	0.615
Contamination mixture	5	0.100	0.119	0.146
SDNN				
IVW	5	−0.005	0.132	0.971
MR Egger	5	0.606	0.484	0.299
Weighted median	5	0.097	0.129	0.454
MR RAPS	5	0.028	0.133	0.832
MR PRESSO	5	−0.005	0.132	0.973
Contamination mixture	5	0.185	0.213	0.177
pvRSA/HF				
IVW	4	0.029	0.045	0.522
MR Egger	4	0.105	0.091	0.367
Weighted median	4	0.047	0.040	0.244
MR RAPS	4	0.037	0.036	0.301
MR PRESSO	4	0.029	0.045	0.568
Contamination mixture	4	0.050	0.109	0.156

Abbreviations: N_{SNPS} = number of single nucleotide polymorphisms, B = beta coefficient, SE = standard error, lCI = lower bound of 95% confidence interval, uCI = upper bound of 95% confidence interval, IVW = inverse variance weighted, MR = Mendelian randomisation, RAPS = Robust Adjusted Profile Score, PRESSO = Pleiotropy RESidual Sum and Outlier. RMSSD = the root mean square of the successive differences of inter-beat intervals; SDNN = the standard deviation of the normal-to-normal inter-beat intervals; pvRSA/HF = the peak-valley respiratory sinus arrhythmia or high-frequency power.

5.3.4 Multivariable MR analysis with resting heart rate and heart rate variability

I also performed multivariable MR using RHR and one HRV measure (RMSSD) as exposures. Variants that were significant in either exposure dataset and available in both datasets were retrieved. These were then clumped and harmonised, resulting in a final set of 18 SNPs. The results from the multivariable MR did not support a causal effect of RHR on ASB when accounting for cardiac vagal effects (Table **5.5**).

Table 5.5 *Results of the multivariable Mendelian randomisation analysis on resting heart rate and heart rate variability on antisocial behaviour.*

Method	N_{SNPs}	*B*	*SE*	*p*
Resting heart rate				
IVW	18	0.016	0.145	0.914
MR Egger †	18	0.096	0.223	0.667
Heart rate variability (RMSSD)				
IVW	18	0.000	0.010	0.992
MR Egger †	18	0.019	0.027	0.482

Note. † estimate when effects are oriented to the respective exposure. Abbreviations: N_{SNPS} = number of single nucleotide polymorphisms, B = beta coefficient, SE = standard error, lCI = lower bound of 95% confidence interval, uCI = upper bound of 95% confidence interval, IVW = inverse variance weighted, MR = Mendelian randomisation, RMSSD = the root mean square of the successive differences of inter-beat intervals.

5.3.5 Univariable MR analysis between resting heart rate and heart rate variability

As a positive control, I conducted an MR analysis using RHR as the exposure and an alternative outcome that I assumed would be causally related to RHR (HRV). These results suggested causal effects between RHR and HRV (N_{SNPs} = 206; B_{IVW} = −0.014; *95% CI* = −0.017, −0.011; $p < 0.001$; Table **5.6**).

Table 5.6 *Results from the positive control analyses using resting heart rate as the exposure and heart rate variability (standard deviation of the normal-to-normal inter-beat intervals; SDNN) as the outcome in a univariable Mendelian randomisation analyses.*

Method	N_{SNPs}	***B***	***SE***	***p***
IVW	206	−0.014	0.002	6.12×10^{-17}
MR Egger	206	−0.021	0.004	1.49×10^{-7}
Weighted median	206	−0.016	0.003	5.64×10^{-8}
MR RAPS	206	−0.014	0.002	2.62×10^{-14}
MR PRESSO	206	−0.014	0.002	7.04×10^{-18}
Contamination mixture	206	−0.024	0.003	8.35×10^{-11}

Abbreviations: N_{SNPS} = number of single nucleotide polymorphisms, B = beta coefficient, SE = standard error, lCI = lower bound of 95% confidence interval, uCI = upper bound of 95% confidence interval, IVW = inverse variance weighted, MR = Mendelian randomisation, RAPS = Robust Adjusted Profile Score, PRESSO = Pleiotropy RESidual Sum and Outlier.

5.3.6 Linkage disequilibrium score regression

Finally, I performed LD score regression to calculate genetic correlations between the heart rate measures and ASB. The results indicated genetic correlations between RHR and HRV (r_{GE} between RHR and HRV (RMSSD) = -0.588, $p = 2.49 \times 10^{-21}$) but no genetic correlations between either heart rate measure and ASB (r_{GE} between RHR and ASB = 0.057, p = 0.169; r_{GE} between HRV (RMSSD) and ASB = 0.164, $p = 0.169$; Table **5.7**).

Table 5.7 *Genetic correlations between resting heart rate, heart rate variability and antisocial behaviour.*

Method	r_g	*SE*	*p*
Resting heart rate			
HRV RMSSD	-0.588	0.062	2.49×10^{-21}
HRV SDNN	-0.560	0.062	2.63×10^{-19}
HRV pvRSA/HF	-0.297	0.057	1.55×10^{-7}
Antisocial behaviour			
Resting heart rate	0.057	0.042	0.169
HRV RMSSD	0.164	0.114	0.153
HRV SDNN	0.024	0.109	0.828
HRV pvRSA/HF	0.122	0.125	0.330

Note. r_g *= Pearson's r for genetic correlation, SE = standard error, p = p-value, HRV = heart rate variability, RMSSD = the root mean square of the successive differences of inter-beat intervals; SDNN = the standard deviation of the normal-to-normal inter-beat intervals; pvRSA/HF = the peak-valley respiratory sinus arrhythmia or high-frequency power.*

5.4 Discussion

In these preliminary analyses using two-sample Mendelian randomisation (MR), I report no causal effects of resting heart rate (RHR) on antisocial behaviour (ASB). Sensitivity analyses suggested that these results are unlikely to be affected by heterogeneity, pleiotropy and/or weak instrument bias. Additional analyses using a measure that captures the vagal contribution to RHR (heart rate variability; HRV) also did not indicate causal effects, and there were no significant genetic correlations between any measure of heart rate and ASB.

In line with two prior studies that controlled for unmeasured confounders (Baker et al., 2009; Kendler et al., 2021) using a twin and a co-relative control design, our results do not support the hypothesis that the often-observed association between RHR and ASB is directly causal. The null findings in the current study and the discrepancy between these and the phenotypic associations reported in previous research lend themselves to several alternative explanations.

First, it may be that the relationship between RHR and ASB *is* causal, but the current study did not have adequate power due to the sample size of the outcome GWAS (n = 85,359). Additionally, it could be that RHR has a causal effect on specific, potentially "more extreme" forms of ASB. The most recent meta-analysis on RHR and ASB found significant evidence of heterogeneity of effects, with the effect of RHR on ASB being the largest for the most violent offenders and psychopathy (de Looff et al., 2022). Although the phenotype used in the current study included "more extreme" forms of ASB (e.g., violent and sexual crimes) and clinical samples, these measures were combined with other "less extreme" forms (e.g., delinquency) to create a broad measure of ASB. Therefore, it is possible that I was not able to detect a potentially true causal effect on specific forms of ASB. However, it should be noted that the only

three other genetically informed studies found no evidence of an effect of RHR on aggression, delinquent behaviours, and psychopathic traits in childhood (Baker et al., 2009) or criminal behaviour, drug abuse and alcohol use disorder in adulthood (Kendler et al., 2021) or a genetic correlation between RHR and childhood aggression (Ip et al., 2021). Future research should investigate heterogeneity in the relationship between RHR and ASB by considering specific phenotypes of ASB.

Another potential explanation is that the relationship between RHR and ASB *is not* causal but may arise partly due to issues in data quality and/or publication bias in previous research. Regarding data quality, many existing studies have included small and/or non-representative samples. In the most recent meta-analysis on RHR and ASB (de Looff et al., 2022), over half of the studies included data from fewer than 100 participants (61%; $k = 62$). The funnel plots also showed evidence of publication bias, whereby extreme negative findings were more likely to have been identified and included in the meta-analysis than studies that reported either a null or positive effect of RHR on ASB. The availability of larger and more representative datasets, such as those included in the current analyses, lends itself to future research overcoming these data quality issues.

A third explanation is that the association reported in previous research is driven by genetic confounding. Indeed, previous genetically informed studies suggest that the association between RHR and ASB is entirely explained by genetic effects (Baker et al., 2009; Kendler et al., 2021). However, using LDSC, I was unable to support this hypothesis. Our results suggested no evidence of genetic correlation between any measure of heart rate and ASB, in line with a recent GWAS on childhood aggression, which also found no genetic correlation with RHR (Ip et al., 2021). It should be noted that LDSC relies on common SNPs, which only capture a fraction of the heritability of RHR and ASB. Therefore, I cannot rule out the possibility that future studies using

larger samples and/or including rarer variants may detect significant genetic correlations between RHR and ASB.

A fourth explanation for the observed association between RHR and ASB is that it is driven by other confounders that were not adequately accounted for in previous research. In another recent meta-analysis (Portnoy & Farrington, 2015), over three-quarters of the effect sizes included did not adjust for any confounders (77%; $k = 89$). It may be, for example, that sensation-seeking behaviour, which is associated with both RHR and ASB (Hammerton et al., 2018; Portnoy et al., 2014; Sijtsema et al., 2010), could be a "common cause" for both low RHR and high ASB. These alternative causes of the association between RHR and ASB must be investigated fully, as simply adjusting for a larger number of putative confounders poses the risk of conditioning on mediators and colliders.

Potential time-varying confounders have also not often been considered previously. Repeated measures enable the examination of the temporal ordering of variables. By using certain causal inference methods, it is possible to control for time-fixed unmeasured confounding (e.g., fixed effects analyses) and/or measured time-varying confounders (e.g., g-methods). However, there is currently a lack of longitudinal studies looking at RHR and ASB. In the same meta-analysis, nine in ten of the studies included were cross-sectional (90%; $k = 91$). Indeed, I am aware of only eight studies with moderate sample sizes (i.e., including more than 100 participants) that use longitudinal data (Baker et al., 2009; Galán et al., 2017; Hammerton et al., 2018; W. G. Jennings et al., 2013; Kavish et al., 2020; Kendler et al., 2021; Latvala et al., 2015; Murray, Hallal, et al., 2016). Five of these studies found an inverse association between RHR and ASB (Baker et al., 2009; Hammerton et al., 2018; W. G. Jennings et al., 2013; Latvala et al., 2015; Murray, Hallal, et al., 2016) and three studies found no relationship (Galán et al., 2017; Kavish et al., 2020; Kendler et al., 2021). This

highlights the inconsistencies in the literature, which may arise in part due to differences in study design, RHR and ASB measurement, analyses, and confounder adjustment. Of note, only two of these studies employed causal inference methods (Baker et al., 2009; Kendler et al., 2021) and both used family-based genetically informed methods. To draw causal conclusions, triangulation using methods that utilise different but complementary assumptions is needed (Burgess et al., 2017; Lawlor et al., 2017). The current study adds to the existing evidence by relying on IV assumptions. However, further analyses with large, longitudinal datasets using causal inference methods are required to disentangle this relationship.

Despite limited confounder adjustment, the use of small sample sizes, cross-sectional data, and evidence of publication bias, it has been argued that RHR measures could be incorporated into risk assessments and interventions for ASB (de Looff et al., 2022; Portnoy & Farrington, 2015). Nevertheless, the null findings from the current study and the lack of high-quality evidence from previous research suggest that, although RHR may still be a robust indicator of ASB, RHR should not be interpreted as a causal risk factor for ASB.

5.4.1 Strengths and limitations

The current study has key strengths, such as utilising data from two large GWAS and using MR analyses, which can help strengthen causal inference when instrumental variable assumptions are met. However, I must consider certain limitations. As mentioned above, the ASB GWAS had a relatively small sample size (n = 85,359) compared to other GWAS and reported an SNP heritability of 8.4% (Tielbeek et al., 2022). Therefore, the current study may not have been powered to detect small causal effects. However, the clinical utility of such small effects, for instance, using RHR as a basis to diagnose or intervene on ASB, is uncertain. As is often the case, once more

data are available, the analyses should be updated using GWAS with larger sample sizes.

Another potential limitation is that the phenotypic measurements in both the exposure and the outcome GWAS were heterogeneous. The exposure GWAS used a measure of RHR averaged over multiple measures, and the outcome GWAS combined questionnaires that captured a broad range of ASB types. Therefore, I may not have been able to detect an association between RHR and ASB due to the heterogeneity in the measurements used.

Furthermore, it should be noted that the exposure GWAS used an older sample than the outcome GWAS sample, meaning that the SNPs-exposure associations were measured later than the SNPs-outcome associations. MR estimates are often interpreted as lifetime exposures, but the effect of this on MR results is an issue of ongoing debate (Sanderson, Glymour, et al., 2022; Sanderson, Richardson, et al., 2022). However, the effects of RHR on ASB may be different over the life course (e.g., during adolescence), and I could not detect this in the current study.

Finally, although it is more of a concern for significant MR findings, some of the instrumental variable assumptions of MR are not verifiable. To be confident that the assumptions were supported, I used genetic variants that reached genome-wide significance, checked for high *F* statistics to support the relevance assumption, ensured the absence of significant horizontal pleiotropy to support the exclusion restriction assumption, and used a positive control design.

5.4.2 Conclusions

In this Chapter, I present results from genetically informed analyses (MR and LDSC) that suggest no significant genetic correlation nor a causal relationship between RHR and ASB. The results do not support the idea that the often-reported association between RHR and ASB is causal. The association reported by observational studies may be due to biased estimates resulting from small, selective samples, publication bias, and/or inadequate control of genetic and environmental confounders. Future research should use larger samples and appropriate controls by using longitudinal data and more robust causal inference methods to understand the relationship between RHR and ASB further.

Key points

1. Associational studies suggest a link between low resting heart rate (RHR) and higher levels of antisocial behaviour (ASB).

2. However, few studies have used causal inference methods, and none have used Mendelian randomisation (MR) to investigate whether this relationship reflects causality.

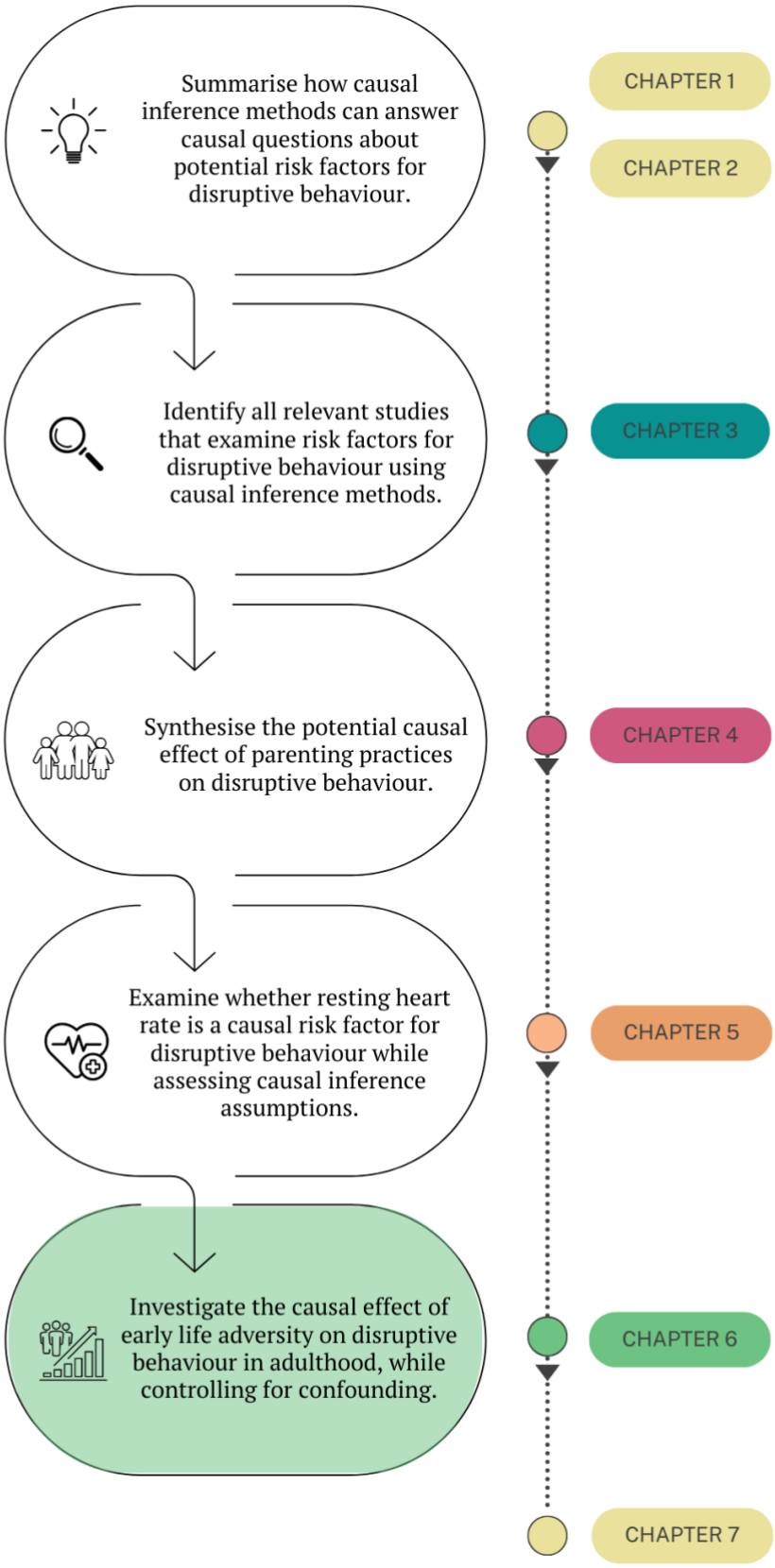
3. In this Chapter, I used two-sample univariable MR, multivariable MR, and linkage disequilibrium score regression (LDSC) analyses to triangulate evidence on the association between RHR and ASB.

4. The findings indicate that RHR does not have a direct causal effect on ASB and that RHR and ASB are not genetically correlated. This suggests that previously observed associations may be biased by confounding, reverse causation, and/or additional study characteristics.

5. Considering the often unverifiable assumptions of MR (i.e. *relevance*, *exchangeability,* and *exclusion restriction*), researchers should aim to instead falsify these assumptions and conduct appropriate sensitivity analyses, as detailed in this Chapter, to quantify the influence of any potential biases.

6. This Chapter demonstrates that MR analyses can provide a quick and efficient way to triangulate evidence on potentially modifiable risk factors for disruptive behaviour.

THESIS STRUCTURE



6 TRAJECTORIES OF EARLY LIFE ADVERSITY AND LATER DISRUPTIVE BEHAVIOR: A NATIONWIDE STUDY OF OVER 1.9 MILLION CHILDREN

Chapter overview

In Chapter 5, I conducted the first Mendelian randomisation analyses on the relationship between resting heart rate and antisocial behaviour. This combination of risk factors and method was informed by my systematic review (Chapter 3), which also revealed that no previous studies have used g-computation to examine risk factors for DBDs. In this Chapter, I will focus on inter-related risk factors for early life adversity that can be assessed through registry-based datasets and have the potential to inform public health initiatives, as many are preventable. Administrative data allows for large sample sizes with long follow-up times and can also address some key biases in observational research, e.g. selection and recall bias.

Publication status: This Chapter is currently in preparation for the *International Journal of Epidemiology:* Karwatowska, L., Viding, E., Pingault., J.-B., Hulvej Rod, N., & De Stavola, B. L. Trajectories of early life adversity and later disruptive behaviour: a nationwide study of over 1.9 million children. (2023).

Note. This Chapter includes preliminary results which may differ slightly from those in subsequent publications.

6.1 Background

Early life adversity, also known as adverse childhood experiences (ACEs), refers to any experiences that occur outside of a child's expected environment, such as poverty, abuse, or severe mental or physical illness within the family (McLaughlin, 2016; The Lancet Public Health, 2021). These experiences are significant enough to require psychological and social adaptation by the individual (E. McCrory et al., 2022; E. McCrory & Viding, 2015). Most research on early life adversity has been conducted in high-income countries (HIC), where it is highly prevalent (Bengtsson et al., 2019; Gilbert et al., 2009; Lai et al., 2019; McLaughlin et al., 2012). Evidence suggests that rates are similar or even higher for some ACEs in lower and middle-income countries (LMIC; Kessler et al., 2010; Krug et al., 2002; Ramiro et al., 2010). Cross-national estimates suggest that 35-80% of individuals experience *at least* one early life adversity (Kessler et al., 2010; McLaughlin et al., 2012; Ramiro et al., 2010). Additionally, individuals who experience one adversity are more likely to experience others (Ben-Shlomo & Kuh, 2002; Dong et al., 2004), as adversities often cluster and co-occur (Briggs et al., 2021; de Vries et al., 2022; Gilbert et al., 2009).

Since the original ACE study (Felitti et al., 1998), other observational studies have repeatedly shown that exposure to early life adversity is associated with negative mental (Baldwin et al., 2023) and physical health outcomes (Bengtsson et al., 2020), including an increased risk of mortality (Rod et al., 2020). ACEs are thought to affect physiological, neurobiological, cognitive, and behavioural pathways, leading to impairments throughout the life-course (Fagundes et al., 2013; E. McCrory et al., 2022; Shonkoff et al., 2012). For instance, individuals who have experienced early life adversity show altered immune function, including increased inflammation (Danese et al., 2007, 2011) and dysregulated hypothalamic-pituitary-adrenal (HPA) axis activity (Heim et al., 2000; Koss & Gunnar, 2018). Furthermore, individuals exposed to early life adversity show altered brain structure and functioning (E. McCrory et al.,

2012) and atypical social-cognitive-affective processing (Berens et al., 2017; E. McCrory & Viding, 2015; McLaughlin et al., 2020). Such sequelae of early adversity are thought to confer "*latent vulnerabilities*" and increase the risk of later negative outcomes (McCrory & Viding, 2015).

For example, not only do individuals who experience early life adversity show differences in their perception, such as perceiving ambiguous social cues as threatening (Keil & Price, 2009), but they also respond differently to stressors, for example, reacting to ambiguous social cues in an aggressive way (Dodge et al., 1995). These latent vulnerabilities have been postulated to affect individuals' socialisation and social relationships over time, resulting in fewer protective factors, e.g. supportive social networks (E. McCrory et al., 2022). These differences affect their ability to make and maintain relationships, which would buffer these effects, thereby contributing to the risk of and maintenance of adverse outcomes (Egeland et al., 2002; Lansford et al., 2007; Maas et al., 2008; Widom & Wilson, 2015). Latent vulnerabilities may be one reason individuals who have experienced adversity in childhood exhibit heightened levels of various mental health disorders, including disruptive behavioural disorders (DBDs; Adjei et al., 2022; Baldwin et al., 2023) and criminal offending (Widom & Wilson, 2015).

Until recently, most evidence on the effects of early life adversity has been limited to results obtained from associational studies, with additional key methodological limitations hindering causal inference in this area. For example, research on adversity has historically been conducted using data from individuals in adulthood (Ceccarelli et al., 2022), which relies on retrospective, subjective measures of adversity that are prone to recall bias (Baldwin et al., 2019, 2023; Francis et al., 2023; Newbury et al., 2018). In addition, much of the research on adversity is cross-sectional; therefore, without information on the temporal ordering of variables, causal inferences on the

relationship between adversity and DBDs cannot be drawn. Further methodological challenges lie in selection bias. For example, individuals who experience early life adversity are less likely to participate in research, and those who do are much more likely to drop out (Doidge et al., 2017). Attrition limits the follow-up period and, consequently, examination of long-term outcomes and can result in collider bias (Munafò et al., 2018).

Previous research has also used varying definitions of adversity. For example, adversities are occasionally studied in isolation. However, failing to account for other co-occurring adversities can confound results as it overly inflates the associations between the adversity of interest being studied and the outcome. To account for this, researchers add multiple adversities together to give a total adversity, or ACE, score (Lacey & Minnis, 2020). A limitation of this approach is that it assumes that different types of adversities confer the same health risks, i.e. of equal severity (Taylor-Robinson et al., 2018). Using an ACE score, a child who lives in poverty and has a parent with a mental health problem would be assigned a score of two, as would a child who experiences both physical and sexual abuse. However, there is evidence that different adversities and combinations of adversities do not carry the same health risks and may work through distinct mechanisms (Lanier et al., 2018; McLaughlin et al., 2020; Putnam et al., 2013).

An alternative approach to dealing with the multi-dimensionality of adversities is group-based multi-trajectory modelling (GBTM; Nagin et al., 2018). GBTM can cluster individuals together based on a set of exposures, such as dimensions of adversity, to investigate whether these combinations are associated with later health outcomes. GBTM has been used to study the effects of adversity using data from cohort studies (Adjei et al., 2022; Lacey et al., 2022; Lanier et al., 2018) and administrative datasets (Bengtsson et al., 2021; El-Khoury et al., 2021; Rod et al., 2020). These studies have

used various combinations of indicators of adversity (e.g. not all include poverty as an indicator of adversity), various periods of exposure (e.g. from birth to early adolescence or birth to adulthood) and, consequently, have derived different numbers of latent classes (between 5 and 7 classes). All the studies found that unique combinations of adversity were associated with various mental and physical health outcomes, including disruptive behaviour (Adjei et al., 2022), post-traumatic stress disorder (El-Khoury et al., 2021), diabetes (Bengtsson et al., 2021), and mortality (Rod et al., 2020).

However, these studies did not examine the effects of adversity in early childhood or estimate the potential impact of intervening on these adversities. Exposure to early life adversity is not inevitable, and every adversity, if truly causal, presents an opportunity for interventions that have the potential to prevent, or at least reduce, the incidence of disruptive behaviour (Gilbert et al., 2009). Adopting a causal inference framework, in addition to supporting aetiological investigations, such as studying the role of combinations of adversities in the aetiology of DBDs, offers methods for addressing questions of the comparative impact of alternative public health interventions (Hernán & Robins, 2016). This is particularly important given the impact of early life in shaping outcomes throughout the life-course (D. J. P. Barker, 1990; E. McCrory et al., 2022).

As such, the current study adds to the literature by examining early life adversity as a potential cause of DBDs. Specifically, I will use GBTM to examine whether individuals can be grouped according to exposure to different dimensions of early life adversity from birth to age 6. To do this, I will use data from an unselected registry-based life course cohort (the DANish LIFE course cohort; DANLIFE). Information on exposures (early life adversity from birth to 6 years) and outcomes (disruptive behaviour from 6 to 25 years) was collected annually on all individuals born in Denmark from 1980 to

2018. This not only addresses issues of selection bias but also recall bias as the measures are prospectively and objectively measured. In line with previous research using DANLIFE, I will use an expert-derived framework to classify adversities across three different dimensions of adversity (material deprivation; loss or threat of loss within the family; negative family dynamics) using GBTM to identify potential aetiological mechanisms of the development of DBDs.

6.2 Methods

6.2.1 Study population

The DANish LIFE Course (DANLIFE) cohort is a prospective register-based study of all children born in Denmark from 1980 until 2015 to mothers who had Civil Personal Registration (CPR) numbers in the Danish Civil Registration System (Bengtsson et al., 2019). A CPR number is a unique 10-digit identification number given to all Danish residents at birth and allows for exact linkage to various nationwide registers that are updated annually. Individuals who live in Denmark but were not born in Denmark (i.e. immigrants) were excluded due to missing information before immigration. The CPR numbers allow family members to be linked. Individuals with missing information on both parents were excluded, as most of the indicators of adversity relied on parental information (n = 3103). The study population includes 2,223,927 children, followed up until 31st December 2017. The total number of years of follow-up differs depending on the year of birth, as illustrated in Figure **6.1**.

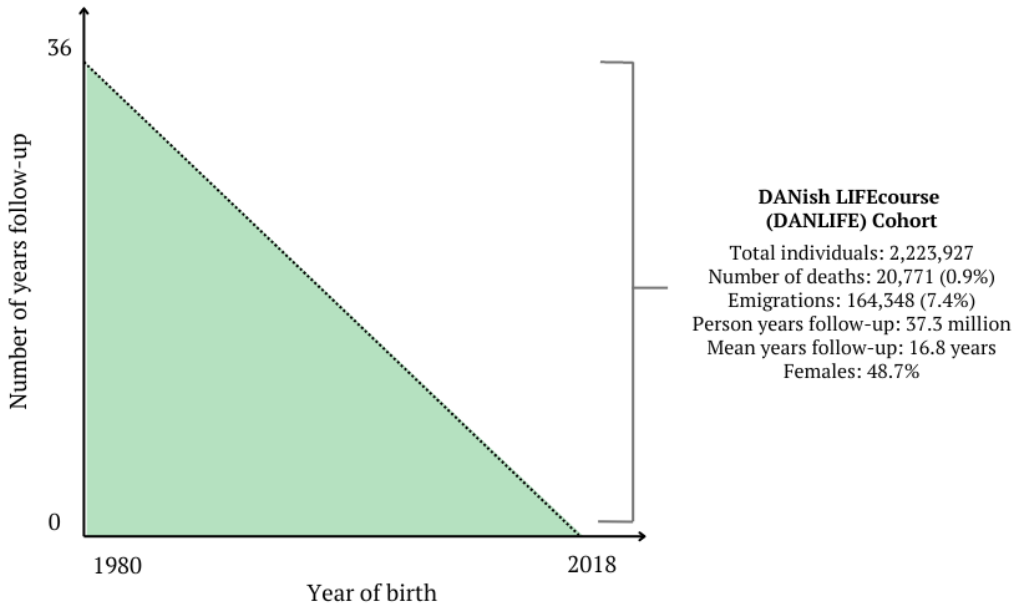


Figure 6.1 *Characteristics of the DANish LIFEcourse (DANLIFE) cohort and follow-up period.*

Data from the most recent version of DANLIFE were accessed on the 24th of August 2023. I excluded individuals who died (*n* = 14,698) or emigrated (*n* = 66,879) before their sixth birthday or who were less than six years old by the end of the follow-up, i.e., were born after 31st December 2011 (*n* = 228,983). I also excluded individuals who were diagnosed with conduct disorder (CD) before their sixth birthday (*n* = 352). Finally, I excluded individuals who did not have valid information on the dimensions contributing to the latent trajectory analyses (*n* = 21,794). The final study population consisted of 1,900,369 individuals (see Figure **6.2**).

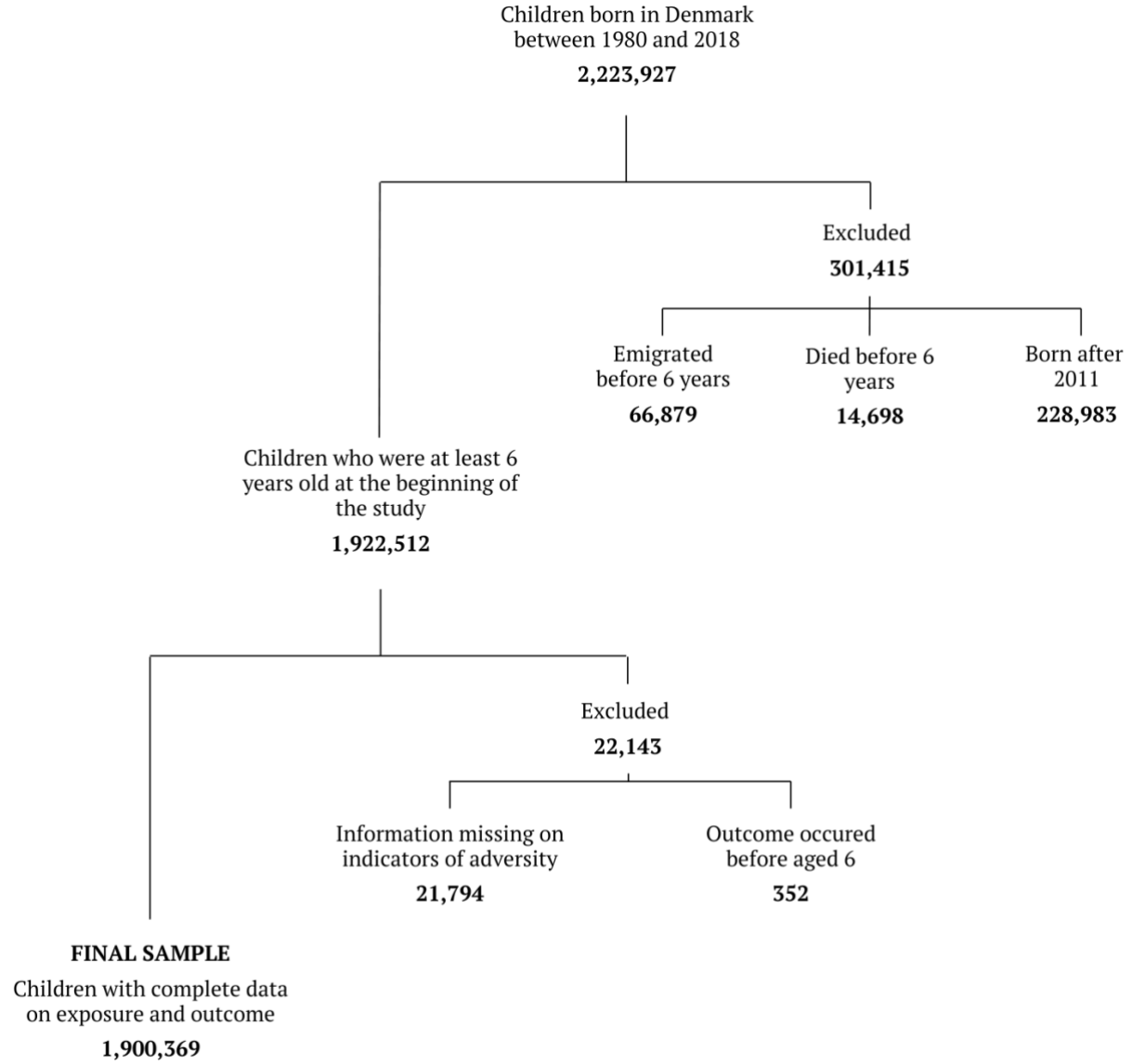


Figure 6.2 *Flowchart of the study population examining the relationship between trajectories of early life adversity and later disruptive behaviour (either conduct disorder diagnoses, dissocial personality disorder diagnoses or convictions of sexual and violent crimes).*

6.2.2 Follow-up

The follow-up period was from the child's sixth birthday until the date of either the outcome of interest (first diagnosis or conviction), emigration, death, their twenty-fifth birthday or the end of follow-up (31st December 2017). Individuals born earlier have longer follow-up periods than those born later (Figure **6.1**). The year that individuals were born will be considered when comparing latent trajectory groups, as clinical and coding practices will likely vary over time. The number of years of follow-up ranged from 0 – 19 years, with a median follow-up time of 16.8 years.

6.2.3 Measures

6.2.3.1 Exposure

The exposure of interest was latent classes of childhood adversity from birth to 6 years. Twelve objectively measured indicators of childhood adversities recorded annually in the Danish registers were grouped into three dimensions by a panel of experts (Rod et al., 2020; Table **6.1**).

These dimensions included:

- **Material deprivation -** expressed in terms of indicators of family poverty and parental long-term unemployment.
- **Loss or threat of loss within the family** - expressed in terms of indicators of parental or sibling severe somatic illnesses or death.
- **Family dynamics -** expressed in terms of indicators of foster care, parental or sibling psychiatric illness, parental alcohol or drug abuse and parental/maternal separation.

From birth to age 6, an individual was categorised as either being exposed or not exposed to a dimension if they experienced any indicator in that dimension. In other

words, each year, an individual could score a maximum of 1 in each dimension and a score of 3 across all dimensions. For example, if, when a child was five years old, their family was in poverty and one of the child's parents was in long-term unemployment, that child would score 1 in the material deprivation dimension at time point 5. If the same child's sibling had a severe somatic illness in the same year, then the index child would also score 1 in the loss or threat of loss within the family dimension. Therefore, at time point 5, the index child would score two across the three dimensions. I analysed these scores using group-based trajectory modelling (GBTM) to identify common trajectories combining adversity dimensions across age (see Section 6.2.4 below).

Table 6.1 *The twelve indicators of adversity, and information sources, used to create the three dimensions of adversity which were included in the latent trajectory analyses.*

	Adversity	Definition	Registry
MATERIAL DEPRIVATION	Family poverty	Family income was below 50% of the median national family income in that specific year.	The Income Statistics Register (Baadsgaard & Quitzau, 2011)
	Parental long-term unemployment	Parent unemployed for at least 12 months within two consecutive years.	The Integrated Database for Labour Market Affiliation (Petersson et al., 2011)
LOSS OR THREAT OF LOSS	Parental severe somatic illnesses	Parent being diagnosed with one of the diseases included in the Charlson comorbidity index (Charlson et al., 1987).	The Danish National Patient Registry (Schmidt et al., 2015)
	Sibling severe somatic illnesses	Sibling diagnosed with one of the seven somatic diagnoses most related to mortality in children aged 0–18 years in Denmark.	The Danish National Patient Registry (Schmidt et al., 2015)
	Parental death	Death of a parent.	The Danish Civil Registration System (Pedersen, 2011)
	Sibling death	Death of a sibling.	The Danish Civil Registration System (Pedersen, 2011)
FAMILY DYNAMICS	Foster care	Placement in out-of-home care.	The Register of Support for Children and Adolescents
	Parental psychiatric illness	Parent admitted for at least one day to a psychiatric hospital or ward with a primary diagnosis related to	The Danish Psychiatric Central Research Register (Mors et al., 2011)

	Adversity	Definition	Registry
		psychiatric illness (excluding primary diagnoses related to alcohol and drug abuse).	The Danish National Patient Registry (Schmidt et al., 2015)
	Sibling psychiatric illness	Sibling being admitted for at least one day to a psychiatric hospital or ward with a primary diagnosis related to psychiatric illness.	The Danish Psychiatric Central Research Register (Mors et al., 2011) The Danish National Patient Registry (Schmidt et al., 2015)
FAMILY DYNAMICS	Parental alcohol use	Parent being diagnosed with a disease related to alcohol abuse or buying a prescribed drug used in treatment of alcohol dependence.	The Danish Psychiatric Central Research Register (Mors et al., 2011) The Danish National Patient Registry (Schmidt et al., 2015) The Danish National Prescription Registry (Wallach Kildemoes et al., 2011)
	Parental drug use	A parent diagnosed with drug dependence or a mental or behavioural disorder due to use of recreational drugs or purchasing a drug prescribed for treatment of drug dependence.	The Danish Psychiatric Central Research Register (Mors et al., 2011) The Danish National Patient Registry (Schmidt et al., 2015) The Danish National Prescription Registry (Wallach Kildemoes et al., 2011)
	Parental/maternal separation	Separation of the parents defined as the parents no longer sharing address.	The Danish Civil Registration System (Pedersen, 2011)

6.2.3.2 Outcomes

Clinical diagnoses of conduct disorder or dissocial personality disorder

The first outcome of interest was diagnoses of either conduct disorder (CD; F91) and/or dissocial personality disorder (DPD; F60) according to the International Classification of Diseases 10th revision (ICD-10; World Health Organization, 2004). Diagnoses were recorded in the Danish Psychiatric Central Research Register from hospital admission records. The ICD-10 defines CD as "characterised by a repetitive and persistent pattern of dissocial, aggressive, or defiant conduct", which constitutes "major violations of age-appropriate social expectations". DPD is defined as

"characterised by disregard for social obligations, and callous unconcern for the feelings of others" and "a low tolerance to frustration and a low threshold for discharge of aggression, including violence."

Convictions of sexual or violent crimes

The second outcome of interest was conviction(s) of either a sexual or violent crime as defined by the Danish Penal Code and registered in the Danish Central Criminal Register. Sexual offences included incest, rape, sexual offences against a child, groping, indecent exposure, prostitution. Violent crimes included violence against public authority, unlawful assembly/disturbance of public order, (attempted) homicide, coercive control, assault, intentional trespass, intentional bodily harm, and causing death or bodily harm by negligence.

6.2.4 Statistical analyses

6.2.4.1 Group-based multi-trajectory models

I used group-based multi-trajectory modelling (GBTM; Nagin et al., 2018), a latent class modelling approach, to identify common clusters of longitudinal indicators of the three prespecified dimensions of adversity (material deprivation; loss or threat of loss within the family; negative family dynamics). I used GBTM instead of alternative latent growth modelling approaches, as it is more flexible than other, more parametrically driven, latent trajectory approaches, such as growth mixture modelling, particularly when dealing with binary latent class indicators because of convergence issues (Herle et al., 2020).

Identification of the best fitting number of latent classes was based on comparing the goodness of fit of models with alternative numbers of classes. The final model was selected based on: *a)* the Bayesian information criterion (BIC), with values closest to zero denoting a better model fit (Nagin et al., 2018); *b)* achieving sample sizes over 5% of the population in each group; *c)* average posterior probabilities of assignment over 0.7; *d)* the odds of correct classification (OCC > 5.0) based on the posterior probabilities of the group.

The TRAJ package in Stata (version 17) was used to fit between two and six trajectory clusters using data from six yearly time points from birth to age six. The trajectories were specified using logistic regressions of each indicator of adversity as a function of age, with separate non-linear terms to capture differences in time trends. Fitted models yielded a probability of each individual being in each cluster (*posterior probabilities*). Individuals were assigned to the group having the highest posterior probability, i.e., using the maximum probability assignment rule.

6.2.4.2 Relationships between the latent classes and outcomes

After identifying the latent classes and assigning individuals to their modal group, I studied the causal associations of these latent adversity groups with the first occurrence of two outcomes (diagnoses and convictions) while accounting for competing events, including death. Previous research using the DANLIFE cohort has shown that all-cause mortality is higher for individuals who experience higher levels of adversity (Rod et al., 2020). The risk of death is also higher for individuals who exhibit disruptive behaviour (Maughan et al., 2014). This introduces the problem of competing events because an individual who experiences death is prevented from experiencing the other outcomes (diagnosis or conviction) despite having increased risks due to their adversity group, thus leading to an underestimate of its effect.

Although mortality rates in this population (< 25 years old) are quite low, mortality is likely to share some factors that also influence diagnoses and convictions and thus induce informative censoring (Hernán et al., 2004; Rod et al., 2020). The two outcomes of interest (diagnoses and convictions) are also likely to share common causes and may, therefore, also act as competing events for each other. Consequently, in the current survival analysis setting, I treat death as a competing event for the two outcomes, diagnoses and convictions, which were considered competing events for each other.

For these reasons, I performed a series of analyses that address the problem of competing events of diagnosis, convictions, and death while targeting the estimation of the relative impact of being in each adversity group versus a relevant reference group (the class with the lowest adversity; Ozenne et al., 2020). These analyses start with the description of the survival rates experienced in this population, separately for each latent trajectory group, to assess the potential competing risks process and then follow a sequence of steps leading to the estimation of population average causal effects (details below).

6.2.4.3 All-cause mortality

To verify whether death is a competing event when considering diagnoses and convictions, I examined whether rates of all-cause mortality varied by latent trajectory group assignment using the Aalen-Johnson estimator to fit cumulative incidence curves and the Cox proportional hazards model (Cox, 1972) to estimate hazard ratios (*HR*) and 95% confidence intervals (95% CI), with and without adjustment for sex and birth year. The proportional hazards assumption, which states that the hazard ratios remain constant over time, was assessed graphically and formally using Schoenfeld's residuals (Schoenfeld, 1982).

6.2.4.4 Cause-specific Cox proportional hazard regression models

I fitted cause-specific Cox proportional hazard regression models (Benichou & Gail, 1990) to estimate the association of the latent trajectory groups and each outcome (diagnoses and convictions) separately, censoring follow-up time when the other outcome, or death, occurred before the outcome of interest, adjusting for birth year and sex. These yielded estimates of conditional cause-specific hazard ratios for the latent trajectories. Note that we have not used other methods for competing events, specifically the Fine and Gray model (Fine & Gray, 1999). I focused on cause-specific Cox modelling as it is a stepping stone towards estimating the average causal effects of the trajectories and is more robust in settings with survival outcomes affected by competing events (Gerds et al., 2023).

6.2.4.5 Average causal effects

Finally, to consider whether there is a causal association between the identified latent trajectory groups and each outcome of interest, I estimated the average treatment effects (ATEs) of being in each latent trajectory group versus the trajectory group with the lowest level of adversity. I express each of these quantities as risk ratios of first diagnosis (or first recorded crime) at pre-selected ages (16, 21 and 25 years) under hypothetical interventions: all individuals in the population being assigned to one of the latent trajectory groups that experienced some form of adversity in childhood versus all individuals being assigned to the adversity trajectory group with the lowest adversity.

Estimation was performed using the g-computation formula, which predicts potential risks at a given age from fitted cause-specific Cox regression models and cumulative incidence curves (Ozenne et al., 2020). Therefore, these risks take into account the

competing events by estimating the cumulative incidence of each outcome *in the presence* of the competing events and are not cause-specific quantities. Confounders are controlled for in the Cox models and then marginalised in the g-computation formula; hence, the predicted potential risks are standardised to the observed (population) distribution of these confounders, making them *marginal* as opposed to conditional on confounders. These risks are then compared, leading to estimates of the ATEs. In other words, the estimates of the ATEs at each age and for each outcome can be interpreted as estimates of population ATEs of each level of exposure versus the reference at the selected age.

The ATEs can be interpreted as causal if the following assumptions are satisfied: *(a) no interference* (i.e. the exposure of one individual does not influence the potential outcome of another person); *(b) counterfactual consistency* (i.e. the exposure groups are well defined); *(c) conditional exchangeability* (i.e. equivalent to no unmeasured confounding); *(d) positivity* (i.e. each individual has a non-zero probability of being in each of the exposure groups); *(e)* correct model specification (Robins, 1986).

6.2.4.6 Covariates

The prevalence of both the exposure (early life adversity) and the outcomes (diagnoses and convictions) are likely to vary by year of birth. Furthermore, parental age at birth may influence the number of adversities experienced before the age of six years and the development of disruptive behaviour. The amount of missing data was higher for paternal age than for maternal age at birth and, as these variables are highly correlated, I chose to only include maternal age at birth. Although other variables associated with social-related or family-related adversity, such as parental education or parental origin, may also affect an individual's adversity trajectory and disruptive behaviour, I did not control for these in the analyses. Parental origin is likely to be

related to disruptive behaviour almost solely through adversity. Furthermore, parental education is likely to be captured by the material deprivation variables. Finally, sex may impact the relationship between exposure to early life adversity and the development of disruptive behaviour. As such, I report all estimates for males and females separately while controlling for birth year and maternal age at birth (see Figure **6.3** for a simplified directed acyclic graph).

Excluding the GBTM, which were fitted using the TRAJ package in Stata (version 17), all other analyses were conducted in R using the "*survival*" (Therneau, 2020) and "*riskRegression*" packages (Gerds et al., 2023).

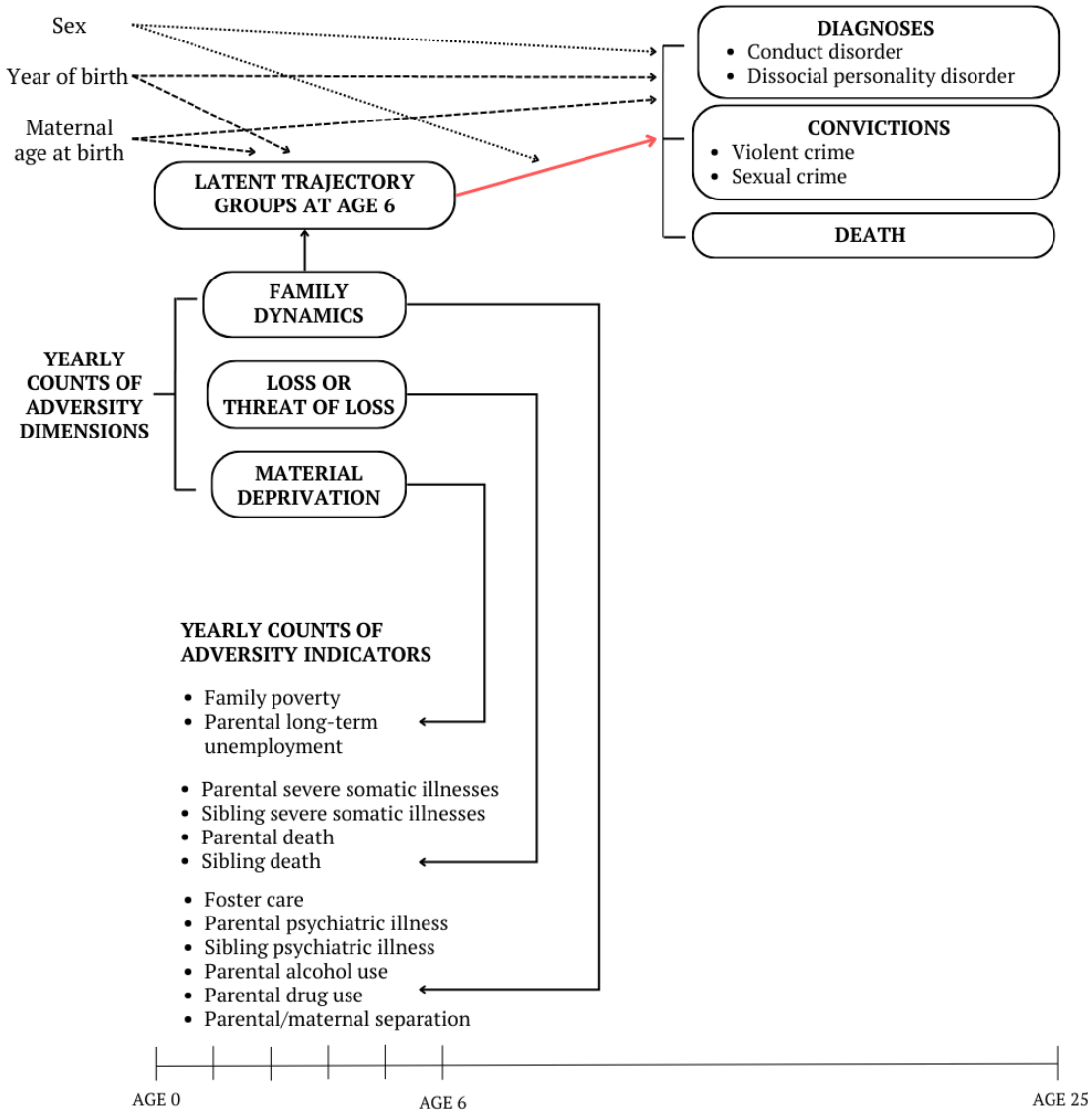


Figure 6.3 *A simplified directed acyclic graph of the assumed pathways between the variables in the analyses.*

6.2.4.7 Ethical permission

Access to the Danish registers is granted by Statistics Denmark and the Danish Health Data Authorities (project reference number: 514–0262/18-3000) and does not require additional ethical permission.

6.3 Results

A total of 1,900,369 children (48.7 % female) from the DANLIFE cohort born between the 1st of January 1980 and the 31st of December 2011 were followed up for 29,217,678 million person-years. The registries recorded 6,502 diagnoses of either CD or DPD (0.3 %) and 35,036 convictions of either a sexual or violent crime (1.8 %) between the ages of 6 and 25 years. Half (49.41%) of individuals in the sample had experienced at least one adversity during the first six years of life (Figure **6.4**). Of the adversities experienced, the most common were parental long-term unemployment (*n* = 430,438; 27.94%), parental separation (301,852; 19.59%) and poverty (192,625; 12.50%; Figure **6.5**).

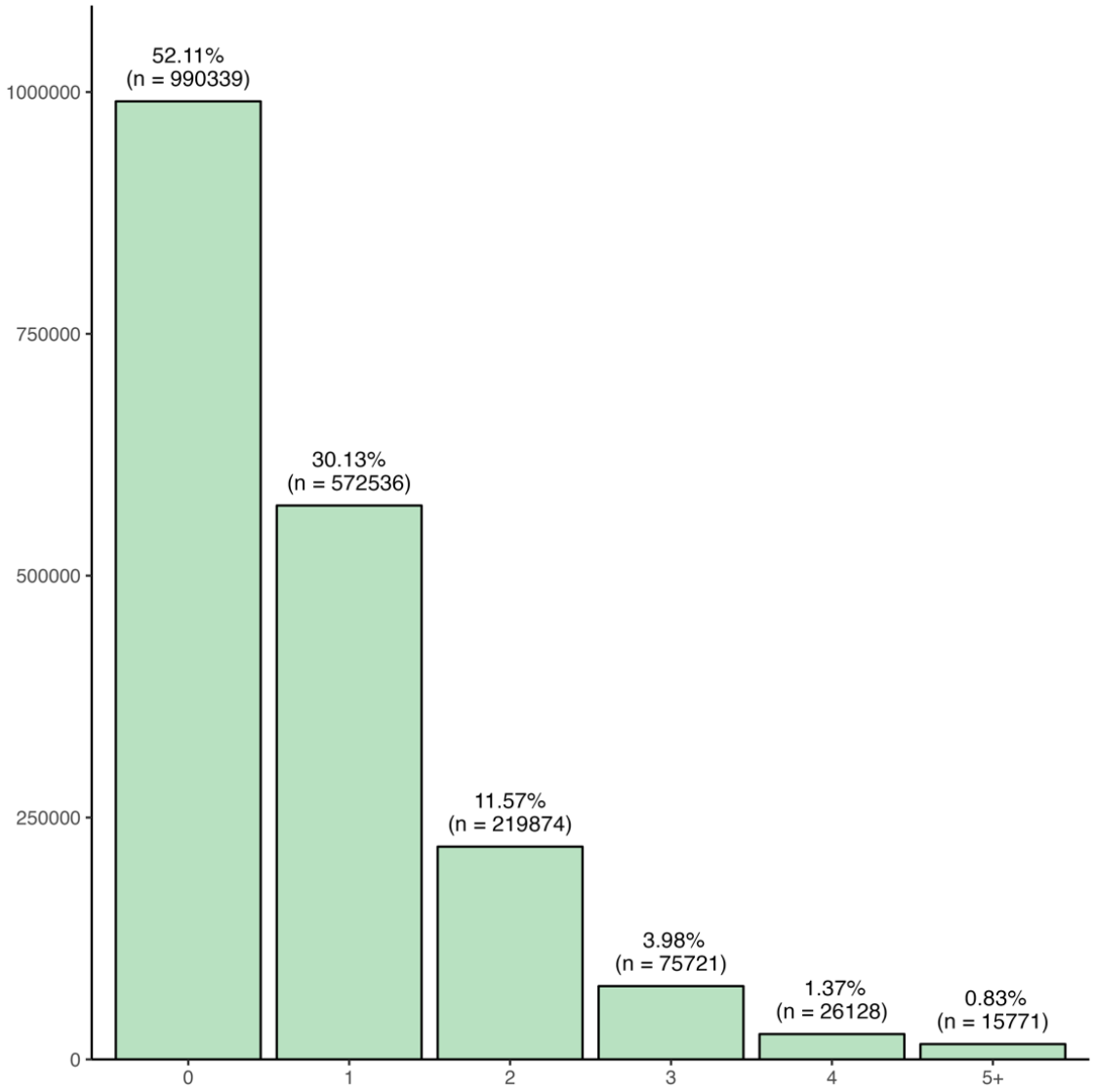


Figure 6.4 *The total number of indicators of adversity that individuals in the DANish LIFEcourse Cohort (DANLIFE) were exposed to from birth to six years.*

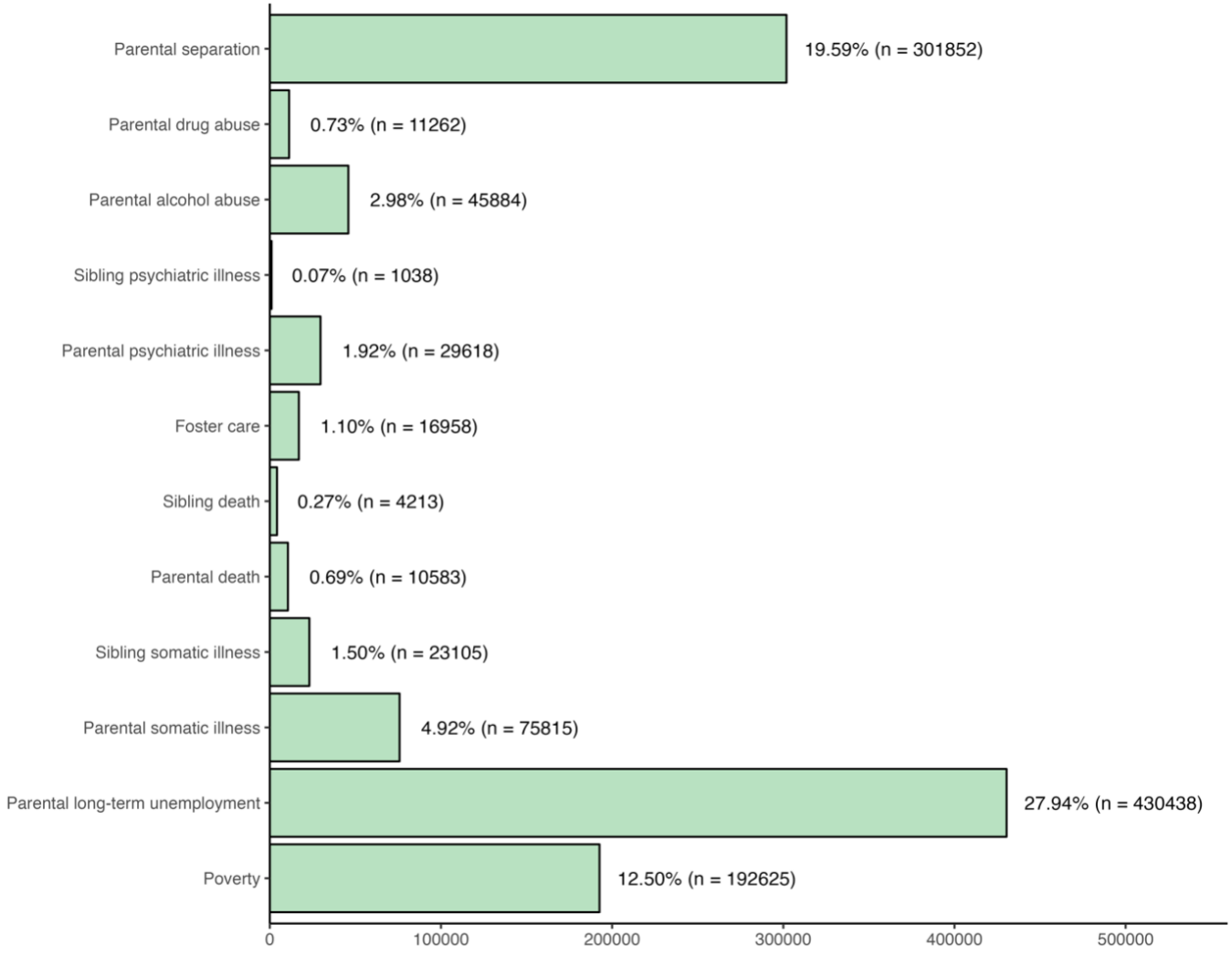


Figure 6.5 *The total number of individuals in the DANish LIFEcourse Cohort (DANLIFE) who were exposed to each indicator of adversity from birth to 6 years old.*

6.3.1 Trajectory groups

I attempted to fit GBTMs with between two and six classes. However, the models with more than four classes did not converge. After examining the model indices, the four-class model was judged to best fit the data as it had the lowest BIC value, each trajectory group had included more than 5% of the population, an average posterior probability of assignment over 0.7 and an odds of correct classification of over 5 (Appendix E, Table 1).

The latent trajectory groups (and their frequencies according to modal assignment) identified were:

- **Low adversity** (LA; *n* = 1,392,151; 68.9%) characterised by a very low annual rate of adversities in all three dimensions before six years of age.
- **Early material deprivation** (early MD; *n* = 226,301; 11.2%) characterised by a high annual rate of material deprivation during the first three years of life followed by a sharp decrease up to 6 years of age.
- **Persistent material deprivation** (persistent MD; *n* = 288,937; 14.3%) characterised by an elevated annual rate of material deprivation during the first three years of life, which increases further up to age six.
- **High loss and negative family dynamics** (high LFD; *n* = 113,151; 5.6%), characterised by a high annual rate of adversity across all three dimensions before six years of age, particularly in the dimensions of loss or threat of loss in the family and negative family dynamics.

The fitted group probabilities for the three dimensions of adversity from birth to age six are shown in Figure **6.6**. The plots for the LA group are shown in blue and indicate low rates throughout the first six years on all three dimensions. Those for the high

LFD group (red) show elevated and increasing levels of loss or threat of loss in the family and negative family dynamics. Those for the early MD group are shown in yellow and have elevated levels in the material deprivation dimension before age three. The persistent MD group (orange) display higher levels of material deprivation, which steadily increase throughout early childhood. The plots for the fitted models with two, three, five and six classes can be found in Appendix E.

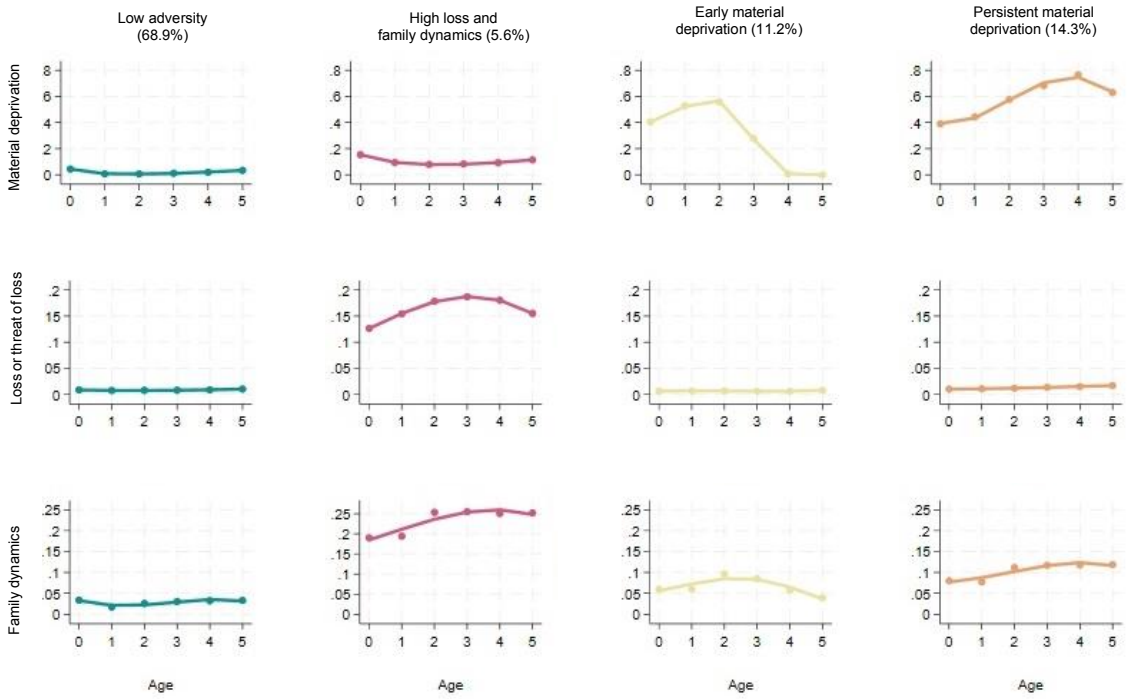


Figure 6.6 *The four trajectories of early life adversity from birth to 6 years old identified by the group-based multi-trajectory models: low adversity (blue); high loss and negative family dynamics (red); early material deprivation (yellow); persistent material deprivation (orange).*

To check whether the latent classes had accurately categorised individuals who had experienced adversity, I plotted the distribution of individuals who had experienced each indicator of adversity from birth to age six by their assigned trajectory group (Figure **6.7**, Figure **6.8**, Figure **6.9**).

Starting with the material deprivation dimension (Figure **6.7**), the largest percentage of individuals who experienced parental long-term unemployment (*n* = 430,438) and poverty (*n* = 192, 625) are those in the early MD and persistent MD groups. The proportions of these two groups are roughly similar until around age three when the percentage of individuals experiencing material deprivation in the early MD group drops dramatically. It is interesting to note that the high LFD group also experiences elevated levels of material deprivation, particularly poverty, from around three years old.

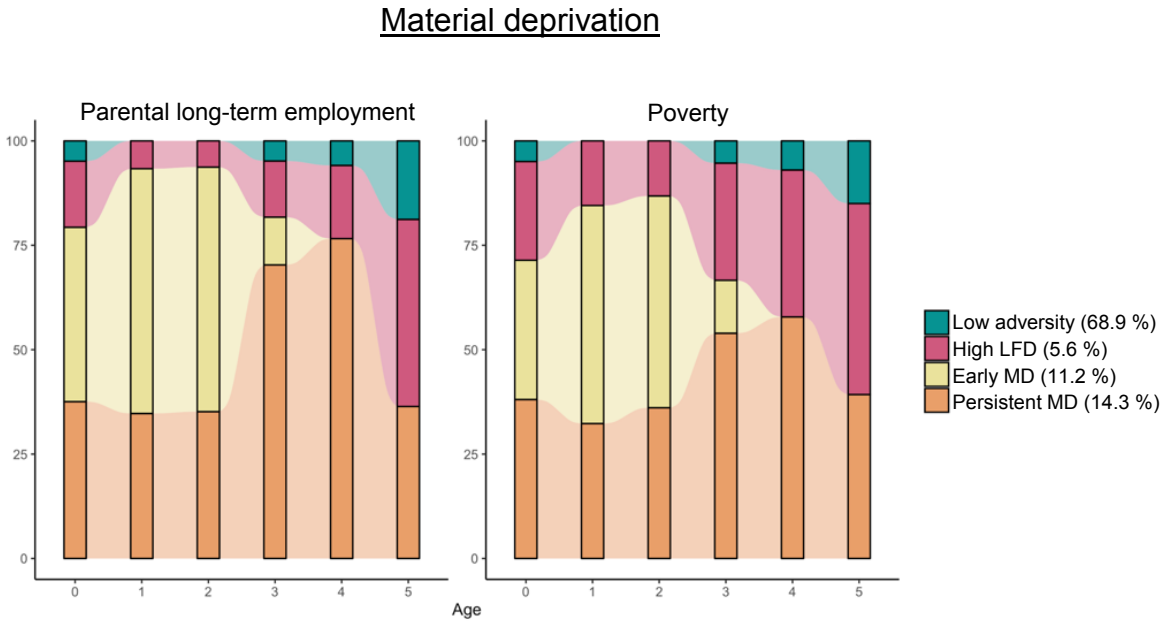


Figure 6.7 *The proportion of individuals exposed to the indicators of adversity for the material deprivation dimension each year from birth to 6 years old by trajectory group.*

In terms of the loss or threat of loss dimension (Figure **6.8**), by far the largest proportion of individuals who experienced any of the four indicators of this adversity (parent severe somatic illness, *n* = 75,815; sibling severe somatic illness, *n* = 23,105; death of a parent, *n* = 10,583; death of a sibling, *n* = 4,213), were from the high LFD trajectory group, which is the smallest trajectory group in terms of numbers of individuals.

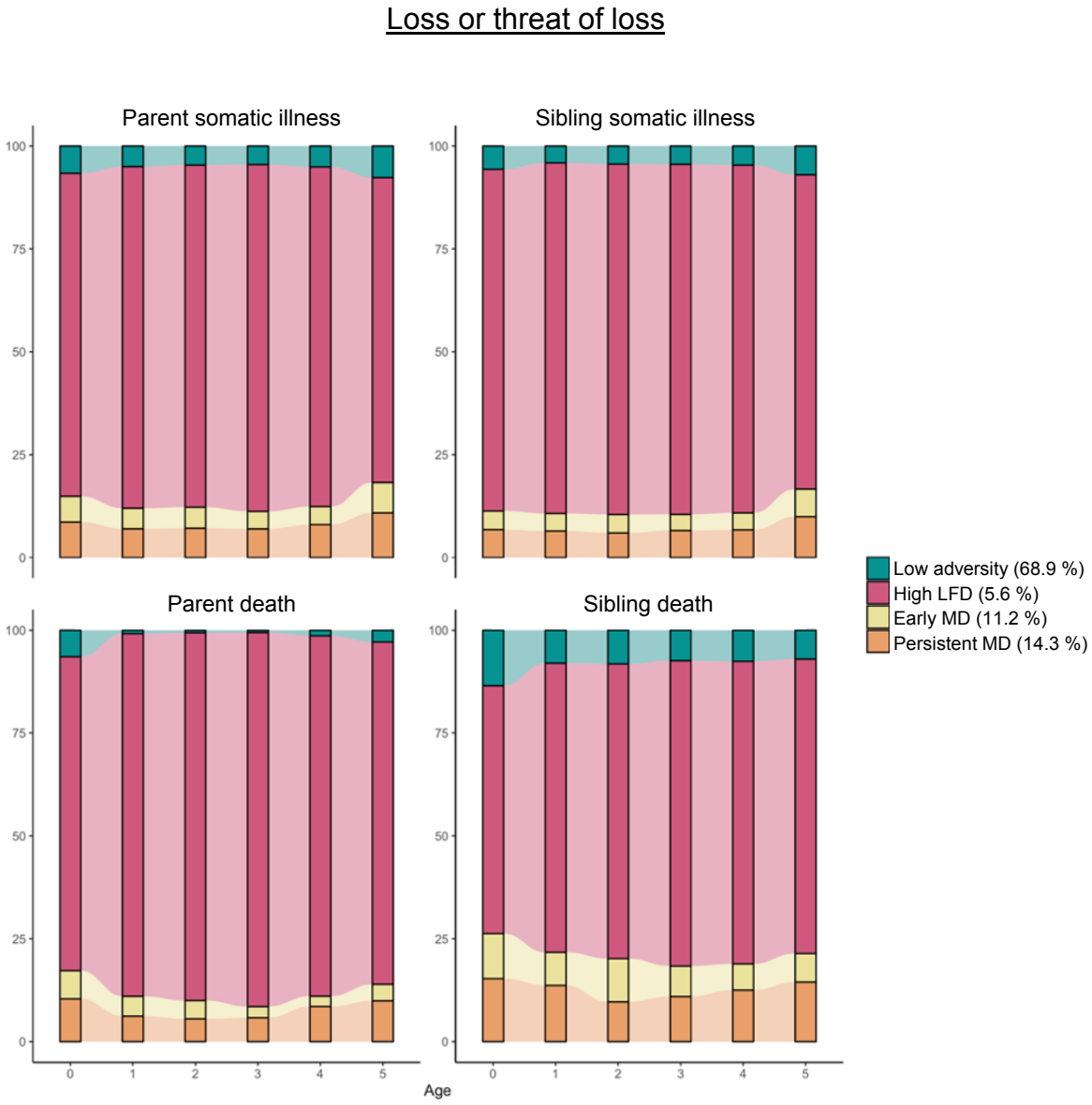


Figure 6.8 *The proportion of individuals exposed to the indicators of adversity for the loss or threat of loss in the family dimension each year from birth to 6 years old by trajectory group.*

Similarly, for each of the six indicators of the family dynamics dimension (parental psychiatric illness, *n* = 29,618; sibling psychiatric illness, *n* = 1,038; foster care, *n* = 16,958; parental separation *n* = 301,852; parental alcohol, *n* = 45,884; and parental drug abuse, *n* = 11,262) the largest percentage of individuals who experienced these adversities were from the high LFD trajectory group, despite this being the smallest trajectory group (Figure **6.9**). The proportion of individuals in each group who experienced parental separation was slightly more even.

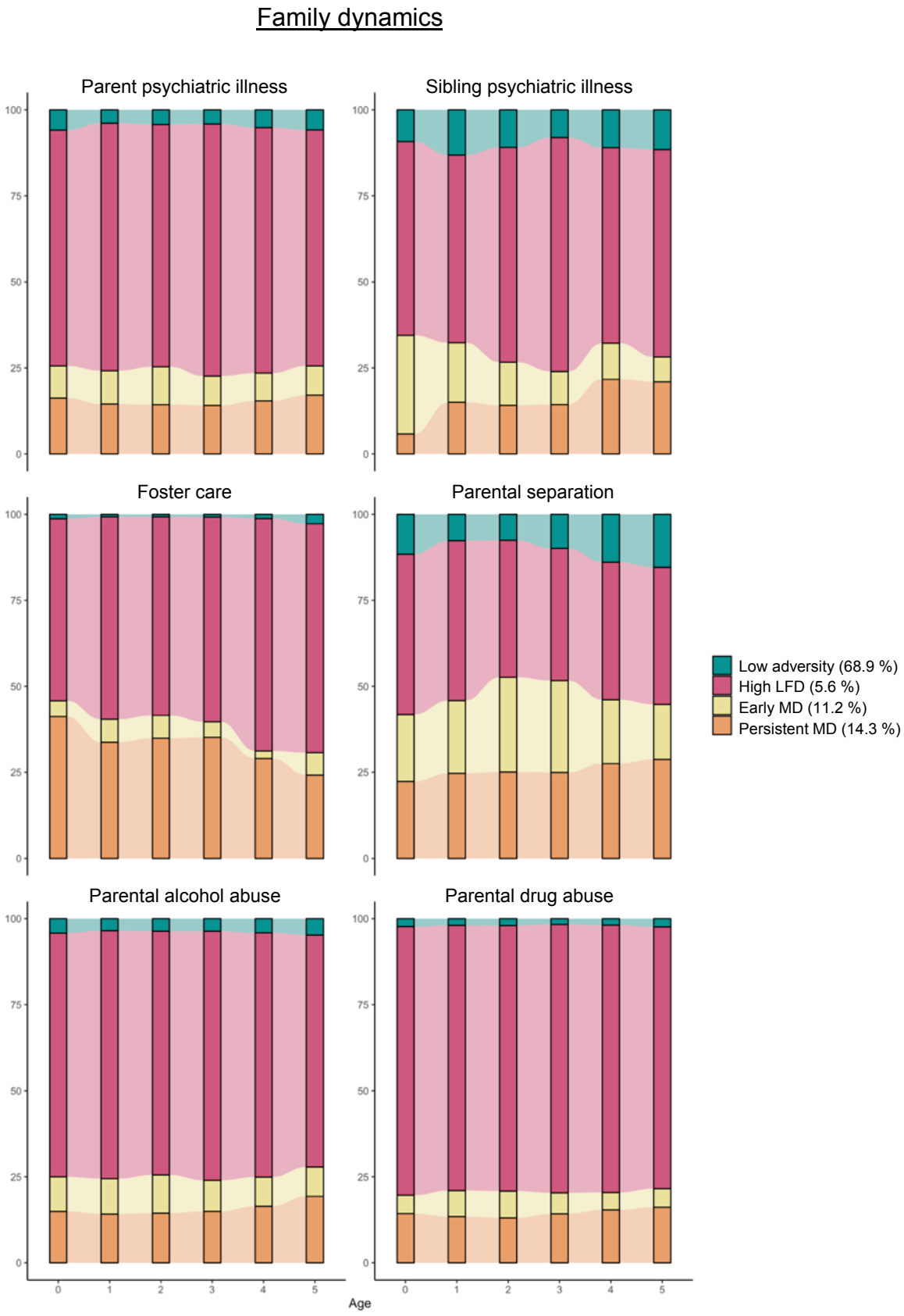


Figure 6.9 *The proportion of individuals exposed to the indicators of adversity for the family dynamics dimension each year from birth to 6 years old by trajectory group.*

6.3.2 Descriptives of the trajectory groups

Figure 6.10 summarises the distribution of key descriptive variables by trajectory group. A similar proportion of females and males were in each of the four groups (see also Table **6.2**). The number of individuals in each group differed according to the year they were born, indicating time trends within the data. A larger proportion of individuals in the two MD groups were born earlier (e.g. between 1980 and 1994), with the proportion of individuals in these groups decreasing in later birth years. On the other hand, the percentage of individuals in the low adversity and high LFD trajectory groups increased, with more individuals in these groups born later. Furthermore, compared to the low adversity group, individuals in the three other trajectory groups were nearly twice as likely to have low birth weight ([low adversity vs highest value for the three other groups] 4.5% versus 7.9%); four times more likely to have low household education (8.4% versus 29.8%); nearly four times more likely to be from a non-European origin (2.4% versus 8.9%); and almost five times more likely to have parents who gave birth under the age of 20 years (maternal age at birth: 1.2% versus 5.5%; paternal age at birth: 0.3% versus 1.4%).

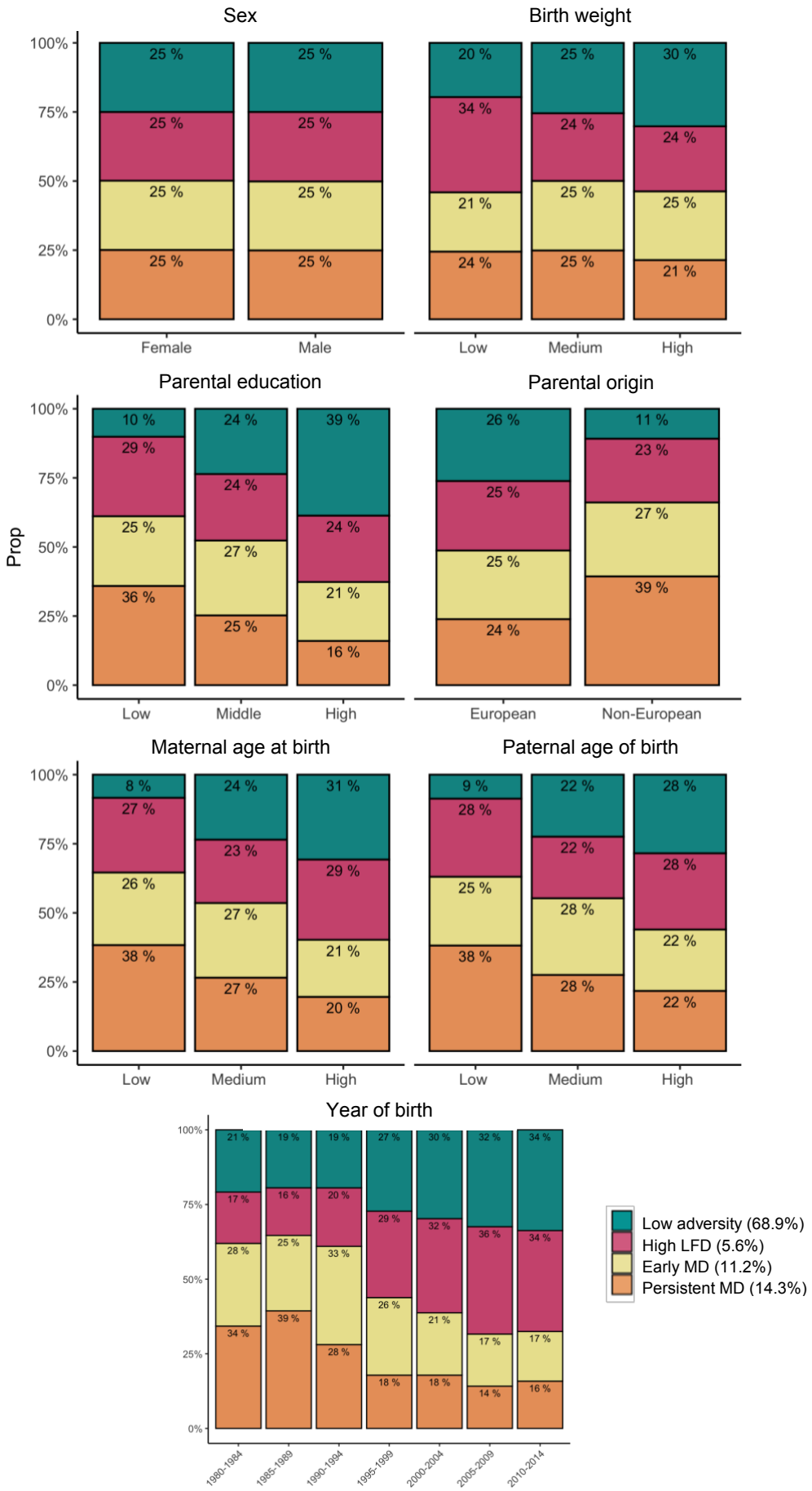


Figure 6.10 *The percentage of individuals in each trajectory group by sex, birth weight category, parental education, parental origin, maternal age at birth, paternal age at birth and year of birth.*

Table 6.2 *Background characteristics of the cohort at the time of birth by the four identified trajectory groups.*

	Overall (*n* = 1,900,369)	**LA** (*n* = 1,371,388)	**High LFD** (*n* = 78,501)	**Early MD** (*n* = 195,516)	**Persistent MD** (*n* = 254,964)
Sex					
Male	974400 (51.3%)	703385 (51.3 %)	40476 (51.6 %)	100150 (51.2 %)	130389 (51.1 %)
Female	925969 (48.7%)	668003 (48.7 %)	38025 (48.4 %)	95366 (48.8 %)	124575 (48.9 %)
Birth year category					
1980-1984	255874 (13.5 %)	165868 (12.1 %)	7833 (10.0 %)	31440 (16.1 %)	50733 (19.9 %)
1985-1989	273831 (14.4 %)	170108 (12.4 %)	8004 (10.2 %)	31570 (16.1 %)	64149 (25.2 %)
1990-1994	319274 (16.8 %)	203618 (14.8 %)	11756 (15.0 %)	49191 (25.2 %)	54709 (21.5 %)
1995-1999	321981 (16.9 %)	244203 (17.8 %)	14850 (18.9 %)	33195 (17.0 %)	29733 (11.7 %)
2000-2004	308936 (16.3 %)	242794 (17.7 %)	14753 (18.8 %)	24293 (12.4 %)	27096 (10.6 %)
2005-2009	304867 (16.0 %)	249639 (18.2 %)	15845 (20.2 %)	19140 (9.8 %)	20243 (7.9 %)
2010-2011	115606 (6.1 %)	95158 (6.9 %)	5460 (7.0 %)	6687 (3.4 %)	8301 (3.3 %)
Birth weight category					
Low	91468 (4.8 %)	61451 (4.5 %)	6180 (7.9 %)	9579 (4.9 %)	14258 (5.6 %)
Medium	1700530 (89.5 %)	1234309 (90.0 %)	67920 (86.5 %)	173971 (89.0 %)	224330 (88.0 %)
High	50427 (2.7 %)	38973 (2.8 %)	1738 (2.2 %)	4577 (2.3 %)	5139 (2.0 %)
Missing	57944 (3.0 %)	36655 (2.7 %)	2663 (3.4 %)	7389 (3.8 %)	11237 (4.4 %)
Highest household education					
Low	251105 (13.2 %)	115328 (8.4 %)	18755 (23.9 %)	40976 (21.0 %)	76046 (29.8 %)
Medium	862603 (45.4 %)	607232 (44.3 %)	35400 (45.1 %)	99349 (50.8 %)	120622 (47.3 %)
High	744194 (39.2 %)	624756 (45.6 %)	22221 (28.3 %)	49160 (25.1 %)	48057 (18.8 %)
Missing	42467 (2.2 %)	24072 (1.8 %)	2125 (2.7 %)	6031 (3.1 %)	10239 (4.0 %)
Parental origin					
Non-European	71935 (3.8 %)	33431 (2.4 %)	4084 (5.2 %)	11794 (6.0 %)	22626 (8.9 %)
European	1793308 (94.4 %)	1318101 (96.1 %)	72581 (92.5 %)	178691 (91.4 %)	223935 (87.8 %)
Missing	35126 (1.8 %)	19856 (1.4 %)	1836 (2.3 %)	5031 (2.6 %)	8403 (3.3 %)

	Overall (*n* = 1,900,369)	LA (*n* = 1,371,388)	High LFD (*n* = 78,501)	Early MD (*n* = 195,516)	Persistent MD (*n* = 254,964)
Maternal age at birth					
<20 years	40785 (2.1 %)	16426 (1.2 %)	3033 (3.9 %)	7342 (3.8 %)	13984 (5.5 %)
20-30 years	1160073 (61.0 %)	811506 (59.2 %)	45272 (57.7 %)	132904 (68 %)	170391 (66.8 %)
> 30 years	664965 (35.0 %)	524117 (38.2 %)	28366 (36.1 %)	50261 (25.7 %)	62221 (24.4 %)
Missing	34546 (1.8 %)	19339 (1.4 %)	1830 (2.3 %)	5009 (2.6 %)	8368 (3.3 %)
Paternal age at birth					
<20 years	10451 (0.5 %)	4330 (0.3 %)	807 (1.0 %)	1771 (0.9 %)	3543 (1.4 %)
20-30 years	844125 (44.4 %)	577583 (42.1 %)	32847 (41.8 %)	101830 (52.1 %)	131865 (51.7 %)
> 30 years	993708 (52.3 %)	758961 (55.3 %)	42142 (53.7 %)	84594 (43.3 %)	108011 (42.4 %)
Missing	52085 (2.7 %)	30514 (2.2 %)	2705 (3.4 %)	7321 (3.7 %)	11545 (4.5 %)

Abbreviations: LA = low adversity; high LFD = high loss and negative family dynamics; early MD = early material deprivation; persistent MD = persistent material deprivation.

6.3.3 Cumulative incidence

Table **6.3** shows the reasons for ending follow-up for individuals in each trajectory group. These include right censoring (either due to emigration or the end of the study), death, diagnoses, or convictions. Table **6.3** also contains information on the number of individuals in each trajectory group who had repeated diagnoses of CD or DPD and/or convictions of sexual and violent crime.

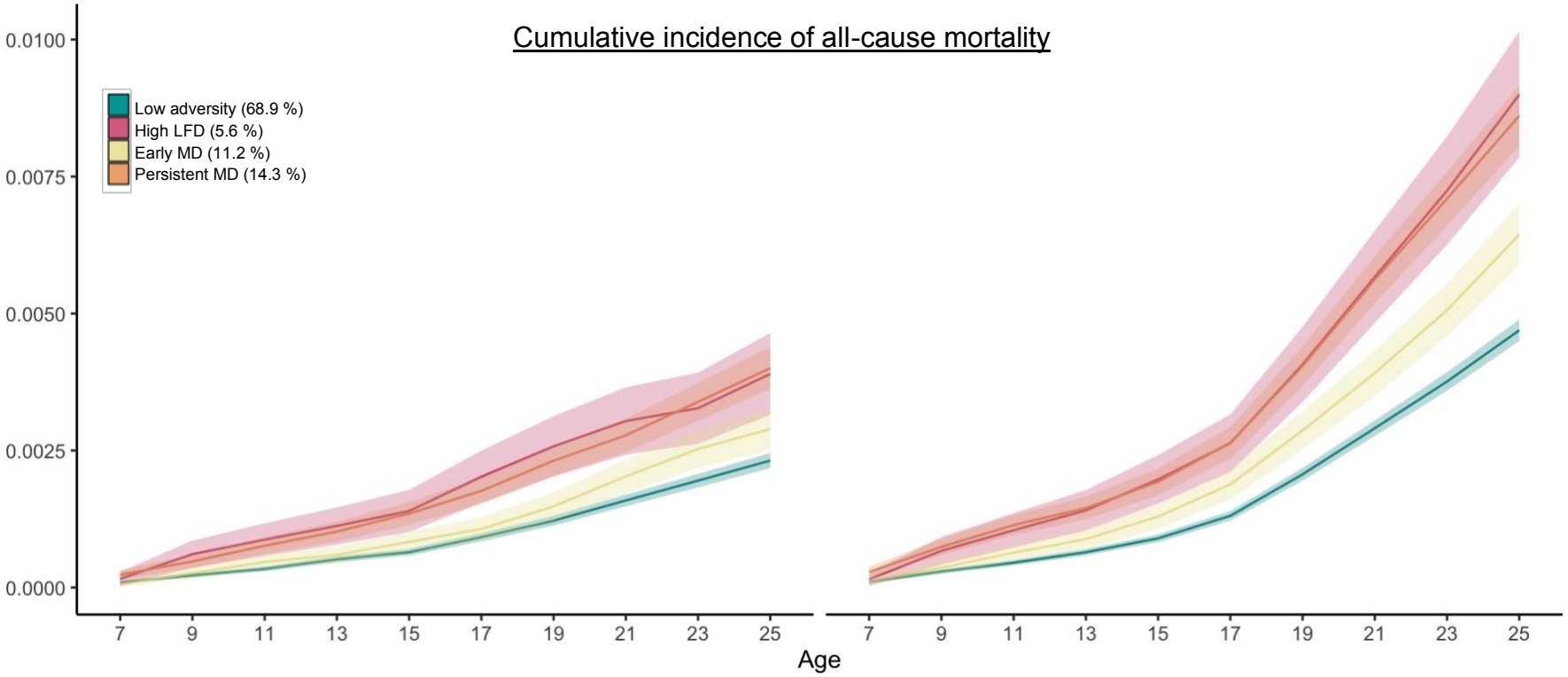
The frequency of death, diagnoses, and convictions by the end of follow-up were all substantially higher in the trajectory groups that experienced early life adversity compared to the LA group. The largest differences were between the LA group and the high LFD group for diagnoses (0.3% vs 0.8% respectively) and the LA group and the persistent MD group for death (0.2% vs 0.5%) and convictions of sexual and violent crimes (1.2% vs 4.1%). This is also reflected in the number of repeated events, with 15% of individuals in the LA group having more than three events recorded versus 22% of individuals in the high LFD and 20% in the persistent MD groups.

Table 6.3 *Number of events and repeated events of the main outcomes (diagnoses of conduct disorder or dissocial personality disorder and convictions of sexual and violent crimes) in each trajectory group.*

	Overall (*n* = 1,900,369)	**LA** (*n* = 1,371,388)	**High LFD** (*n* = 78,501)	**Early MD** (*n* = 195,516)	**Persistent MD** (*n* = 254,964)
End of follow-up					
Right censor	1853317	1347466	75241	188629	241981
	(97.5 %)	(98.3 %)	(95.8 %)	(96.5 %)	(94.9 %)
Death	5514	3256	325	692	1241
	(0.3 %)	(0.2 %)	(0.4 %)	(0.4 %)	(0.5 %)
Diagnoses	6502	3616	601	875	1410
	(0.3 %)	(0.3 %)	(0.8 %)	(0.4 %)	(0.6 %)
Convictions	35036	17050	2334	5320	10332
	(1.8 %)	(1.2 %)	(3.0 %)	(2.7 %)	(4.1 %)
Number of diagnoses and convictions					
1	32284	16675	2053	4741	8815
	(64.4%)	(67.6%)	(59.8%)	(63.6%)	(60.4%)
2	9265	4279	643	1433	2910
	(18.5%)	(17.4%)	(18.7%)	(19.2%)	(20.0%)
3+	8603	3707	736	1278	2882
	(17.1%)	(15.0%)	(21.5%)	(17.2%)	(19.7%)

Abbreviations: LA = low adversity; high LFD = high loss and negative family dynamics; early MD = early material deprivation; persistent MD = persistent material deprivation. Right censor DESCRIPTION

The cumulative incidence rates of all-cause mortality (Figure **6.11**), diagnoses of CD and DPD (Figure **6.12**) and convictions of sexual and violent crimes (Figure **6.13**) estimated using the Aalen-Johnson method are shown in the figures below, separately for males (right) and females (left). All HRs reported in the figure tables are adjusted for birth year. The HRs are estimated from a standard Cox proportional hazard model for all-cause mortality. Meanwhile, the HRs for diagnoses and convictions are estimated from the *cause-specific* Cox models, which I report in Section 6.3.4 below. There is a noticeable difference in the cumulative incidence rates between trajectory groups for all three events. Cumulative incidence rates were highest in the high LFD group, followed by the persistent MD group and the early MD group relative to the LA group. This pattern was evident for all-cause mortality but was even more pronounced for diagnoses of CD and DPD, and the largest differences can be seen for convictions of sexual and violent crimes.



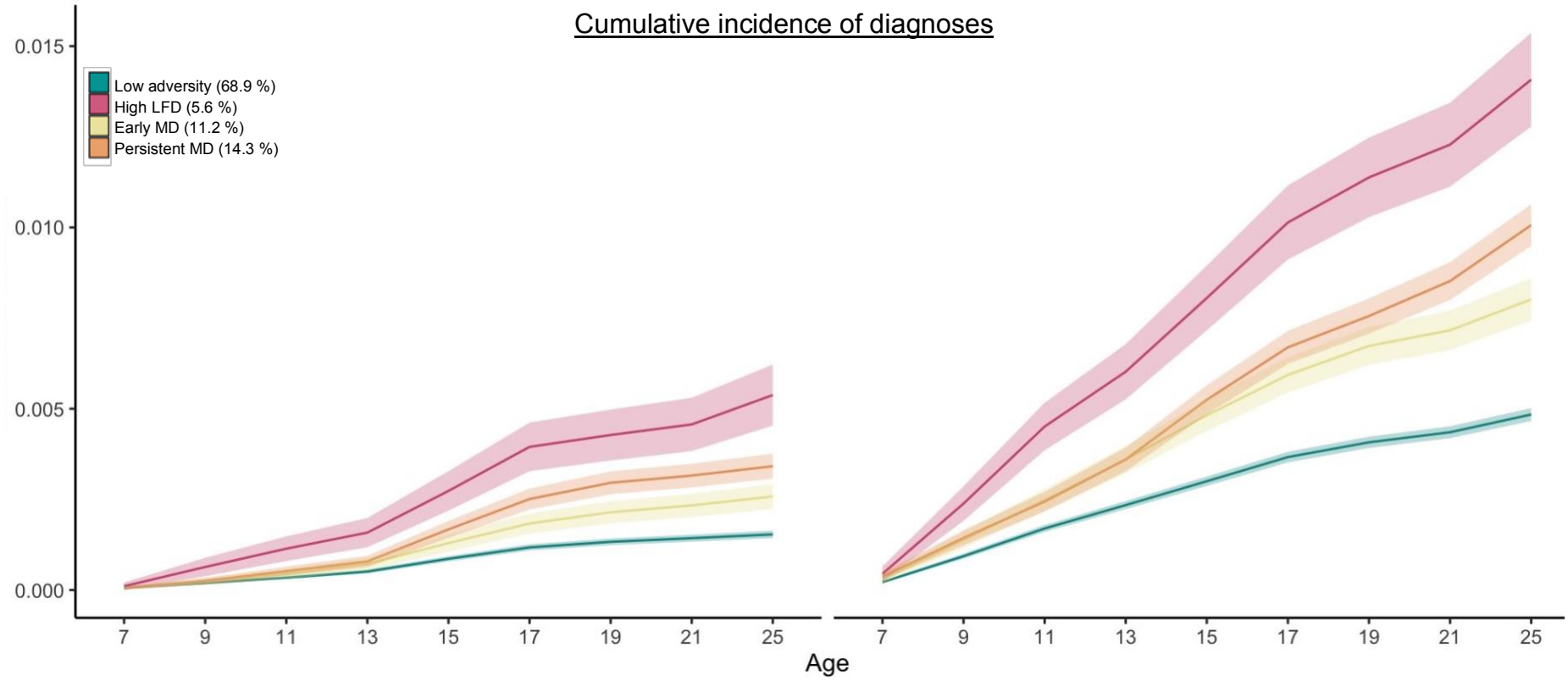
Females

	Total *n*	Deaths	Hazard ratio (*95% CI*)
Low adversity	668,003	1124	REF
High LFD	38,025	112	1.81 (1.49, 2.21)
Early MD	95,366	228	1.22 (1.06, 1.41)
Persistent MD	124,575	425	1.71 (1.53, 1.92)

Males

	Total *n*	Deaths	Hazard ratio (*95% CI*)
Low adversity	703,385	2290	REF
High LFD	40,476	248	1.94 (1.70, 2.22)
Early MD	100,150	520	1.35 (1.22, 1.48)
Persistent MD	130,389	941	1.81 (1.68, 1.96)

Figure 6.11 *Cumulative incidence and birth year adjusted hazard ratios (95% Cis) up to age 25 among 1,900,369 Danish children.*



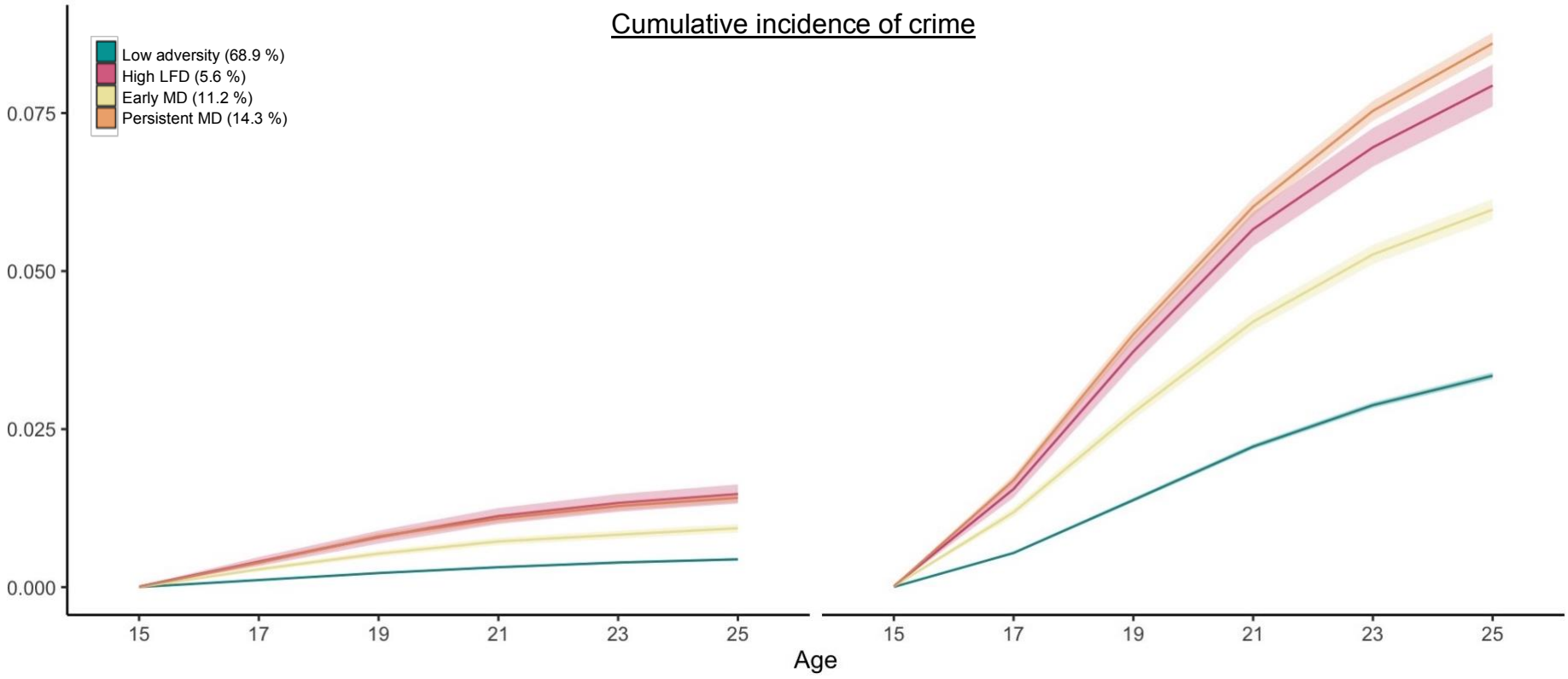
Females

	Total *n*	Diagnoses	Hazard ratio (*95% CI*)
Low adversity	668,003	866	REF
High LFD	38,025	165	3.44 (2.91, 4.07)
Early MD	95,366	219	1.61 (1.38, 1.87)
Persistent MD	124,575	382	2.15 (1.90, 2.43)

Males

	Total *n*	Diagnoses	Hazard ratio (*95% CI*)
Low adversity	703,385	2,929	REF
High LFD	40,476	474	2.80 (2.53, 3.10)
Early MD	100,150	729	1.61 (1.48, 1.75)
Persistent MD	130,389	1,176	1.94 (1.80, 2.08)

Figure 6.12 *Cumulative incidence and birth year adjusted hazard ratios (95% Cis) of diagnoses of conduct disorder or dissocial personality disorder up to age 25 among 1,900,369 Danish children.*



Females

	Total *n*	Convictions	Hazard ratio (*95% CI*)
Low adversity	668,003	1,990	REF
High LFD	38,025	373	3.32 (2.96, 3.72)
Early MD	95,366	718	2.17 (1.99, 2.37)
Persistent MD	124,575	1,455	3.29 (3.07, 3.52)

Males

	Total *n*	Convictions	Hazard ratio (*95% CI*)
Low adversity	703,385	15,561	REF
High LFD	40,476	2,085	2.44 (2.33, 2.56)
Early MD	100,150	4,773	1.83 (1.77, 1.89)
Persistent MD	130,389	9,190	2.68 (2.61, 2.75)

Figure 6.13 *Cumulative incidence and birth year adjusted hazard ratios (95% Cis) of convictions of violent or sexual crimes up to age 25 among 1,900,369 Danish children.*

6.3.4 Cause-specific Cox models

The adjusted hazard ratios estimated from cause-specific Cox models are reported in Table **6.4**. They should be interpreted as conditional associations between the latent trajectory groups with each outcome *in the absence* of the competing events. The proportional hazards assumption was not met for sex (Appendix E), and thus, the results are reported separately for males and females. Moreover, in males only, there was an indication that the proportional hazards assumption was violated by birth year and trajectory groups. However, the smoothed Schoenfeld residuals were approximately horizontal over time, and thus, the trajectory-specific HRs are interpreted as averages of possibly time-varying effects.

Table 6.4 *Estimated hazard ratio and 95% confidence intervals (HR [95% CI]) from cause-specific Cox models of early life adversity and diagnoses of conduct disorder and dissocial personality disorder.*

	Diagnoses		**Convictions**	
	Unadjusted	**Adjusted**	**Unadjusted**	**Adjusted**
	HR (95% CI)	*HR (95% CI)*	*HR (95% CI)*	*HR (95% CI)*
Females				
LA	1	1	1	1
High LFD	3.44 (2.91, 4.07)	3.17 (2.67, 3.77)	3.32 (2.96, 3.72)	3.12 (2.77, 3.52)
Early MD	1.61 (1.38, 1.87)	1.58 (1.35, 1.85)	2.17 (1.99, 2.37)	1.91 (1.75, 2.10)
Persistent MD	2.15 (1.90, 2.43)	2.12 (1.86, 2.42)	3.29 (3.07, 3.52)	3.02 (2.80, 3.25)
Males				
LA	1		1	
High LFD	2.80 (2.53, 3.10)	2.65 (2.39, 2.93)	2.44 (2.33, 2.56)	2.39 (2.27, 2.50)
Early MD	1.61 (1.48, 1.75)	1.57 (1.43, 1.71)	1.83 (1.77, 1.89)	1.66 (1.60, 1.71)
Persistent MD	1.94 (1.80, 2.08)	1.91 (1.77, 2.06)	2.68 (2.61, 2.75)	2.30 (2.24, 2.37)

Note. adjusted models include year of birth and maternal age at birth; Abbreviations: LA = low adversity; high LFD = high loss and negative family dynamics; early MD = early material deprivation; persistent MD = persistent material deprivation.

6.3.4.1 Diagnoses of conduct disorder and dissocial personality disorder

Rates for diagnoses of CD and DPD and convictions of sexual and violent crimes were all higher in the three trajectory groups that experienced adversity compared to the

LA group, with males and females experiencing similar patterns of association. There was a slight decrease in the estimates across all groups after controlling for birth year and maternal age at birth, apart from the risk of diagnoses for early persistent MD, which increased slightly for both males and females (see Table **6.4**).

Compared to individuals in the LA trajectory group, males who experienced high LFD had 2.7 times the rate of diagnoses (adjusted hazard ratio [*a-HR*] = 2.65; *95% CI* = 2.39, 2.93), and women had 3.2 times the rate of diagnoses (*a-HR* = 3.17; *95% CI* = 2.67, 3.77) in the absence of the competing events. Males and females who experienced persistent MD had twice the rate of diagnoses than individuals who experienced low adversity (*a-HR* for males = 1.91, *95% CI* = 1.77, 2.06; *a-HR* for females = 2.12, *95% CI* =1.86, 2.42). Individuals who experienced early MD, compared to those who experienced low adversity, had 1.6 times the rate of diagnoses (*a-HR* for males: 1.57, *95% CI* = 1.77, 2.06; *a-HR* for females: 1.58, *95% CI* = 1.35, 1.85).

6.3.4.2 Convictions of sexual and violent crimes

Compared to individuals in the LA group, males in the high LFD group had 2.4 times the rate of convictions (*a-HR* = 2.39; *95% CI* = 2.27, 2.50), and females had 3.1 times the rate of convictions (*a-HR* = 3.12; *95% CI* = 2.77, 3.52) of sexual and violent crimes, in the absence of competing events. Individuals who experienced persistent MD in early childhood also showed high rates of convictions when compared to those who experienced low adversity, with males and females in the persistent MD group having 2.3 times higher (*a-HR* = 2.24; *95% CI* = 2.24, 2.37) and 3 times higher (*a-HR* = 3.02; *95% CI* = 2.80, 3.25) rates of convictions by the age of 25 respectively. Early MD also showed an association with the rate of convictions, with males showing 1.7 times higher (*a-HR* = 1.66; *95% CI* = 1.60, 1.71) and women showing 1.9 times higher (*a-HR* = 1.91; *95% CI* = 1.75, 2.10) rate of convictions when compared to individuals in the LA group.

6.3.5 Average treatment effects

I predicted the potential risks of each outcome under different "hypothetical" interventions at the age of 6 years, e.g. comparing the potential outcomes of assigning the entire population to the LA trajectory group versus assigning the entire population to the high LFD trajectory group. ATEs comparing each trajectory group that experienced adversity to the LA group were expressed as the ratio of the mean predicted probabilities of each outcome (diagnoses or convictions) by pre-selected developmentally relevant ages (aged 16, 21 and 25 years) under these two interventions. The results across these ages indicated that the relative effects of these hypothetical interventions stayed constant for both outcomes (Table **6.5**).

Table 6.5 *The average treatment effects (expressed as risk ratios [RR]) on diagnoses of conduct disorder and dissocial personality disorder and on convictions of sexual and violent crime of being assigned to one of the adversity groups that experienced adversity compared to the low adversity group, adjusted for year of birth and maternal age at birth.*

	Diagnoses *RR (95% CI)*	**Convictions** *RR (95% CI)*
Females		
High LFD vs. LA		
16 years	3.13 (2.60, 3.69)	3.03 (2.72, 3.41)
21 years	3.12 (2.60, 3.69)	3.02 (2.71, 3.40)
25 years	3.12 (2.59, 3.68)	3.02 (2.71, 3.39)
Early MD vs. LA		
16 years	1.54 (1.29, 1.79)	1.96 (1.79, 2.12)
21 years	1.54 (1.29, 1.79)	1.95 (1.79, 2.12)
25 years	1.54 (1.29, 1.79)	1.95 (1.79, 2.12)
Persistent MD vs. LA		
16 years	2.09 (1.77, 2.34)	2.87 (2.66, 3.10)
21 years	2.08 (1.77, 2.34)	2.86 (2.65, 3.09)
25 years	2.08 (1.77, 2.34)	2.86 (2.65, 3.08)
Males		
High LFD vs. LA		
16 years	2.55 (2.28, 2.82)	2.30 (2.20, 2.41)
21 years	2.54 (2.27, 2.81)	2.27 (2.17, 2.37)
25 years	2.54 (2.27, 2.80)	2.25 (2.15, 2.35)
Early MD vs. LA		
16 years	1.56 (1.44, 1.7)	1.59 (1.53, 1.66)
21 years	1.55 (1.43, 1.7)	1.58 (1.52, 1.64)
25 years	1.55 (1.43, 1.7)	1.58 (1.52, 1.64)
Persistent MD vs. LA		
16 years	1.88 (1.76, 2.00)	2.18 (2.13, 2.25)
21 years	1.88 (1.76, 2.00)	2.15 (2.10, 2.21)
25 years	1.87 (1.75, 1.99)	2.14 (2.08, 2.20)

Note. All models control for birth year. Abbreviations: RR = risk ratio, CI = 95% confidence intervals, LA = low adversity, high LFD = high loss and negative family dynamics, early MD = early material deprivation, persistent MD = persistent material deprivation.

The ATEs on both outcomes were largest for males and females in the high LFD group. Comparing the risks if the entire population was assigned to the high LFD versus to the LA group, the ATE for diagnoses of CD and DPD at age 25 was 2.54 (*95% CI* = 2.27, 2.80) for males and 3.12 (*95% CI* = 2.59, 3.68) for females, which implies that the

probability of diagnoses of CD and DPD was 2.5 times lower for men and 3.1 times lower for women when assigned to the LA group. The *RR* estimates in this group were similar but slightly lower for convictions of sexual and violent crime; by age 25, the estimated ATE was 2.25 (*95% CI* = 2.15, 2.35) for males and 3.02 (*95% CI* = 2.71, 3.39) times lower for females when assigned to the LA group.

The second largest ATEs were observed when comparing the risks if the entire population was assigned to persistent MD compared to the LA group; the probability of diagnosis by the age of 25 was 1.9 times lower for males (*RR* = 1.87; *95% CI* = 1.75, 1.99) and 2.1 times lower for females (*RR* = 2.08; *95% CI* = 1.77, 2.34) in the persistent MD when they were assigned to LA group. In terms of convictions, males exposed to persistent MD had a 2.1 times lower probability (*RR* = 2.14; *95% CI* = 2.08, 2.20), and females had a 2.9 times lower probability (*RR* = 2.86; *95% CI* = 2.65, 3.08) of having a conviction by the age of 25 when assigned to the LA group.

Finally, the lowest ATEs was found in the early MD group; both males and females in this group had a 1.5 times lower probability of diagnoses at age 25 when assigned to the LA group (males: *RR* = 1.55, *95% CI* = 1.43, 1.70; females: *RR* = 1.54, *95% CI* = 1.29, 1.79). Similarly, for convictions of sexual and violent crimes, males in the early MD group had a 1.6 times lower probability (*RR* = 1.58; *95% CI* = 1.52, 1.64), and women had a 2 times lower probability (*RR* = 1.95; *95% CI* = 1.79, 2.12) at the age of 25 when assigned to the LA group.

6.4 Discussion

In this Chapter, I analysed data from an unselected sample of over 1.9 million children from the DANish LIFEcourse cohort (DANLIFE). In line with previous research on DANLIFE, I considered three dimensions of adversity (material deprivation, loss, or threat of loss of a family member and negative family dynamics). Almost half of the sample experienced at least one indicator of these adversities before the age of 6 years. The results from the group-based multi-trajectory modelling (GBTM; Nagin et al., 2018) identified the presence of four latent trajectory groups of adversity which categorised individuals depending on whether they experienced low adversity across all three dimensions ("low adversity"; LA), elevated levels on the loss or threat of loss and negative family dynamics dimensions ("high loss and family dynamics"; high LFD); elevated levels of material deprivation in the first few years of life which reduced over time ("early material deprivation"; early MD); elevated levels of material deprivation throughout childhood which increased from birth to age six ("persistent material deprivation"; persistent MD).

The results from the cause-specific Cox regression models suggested that individuals who experienced any form of adversity experienced higher rates for all outcomes than the LA group. Individuals exposed to the highest levels of adversity (i.e. those in the high LFD group) showed the highest rates of the outcomes, followed by those who experienced persistent MD and then those who experienced early MD, conditionally on year of birth and in the absence of competing events. I predicted the potential risks of each outcome under different "hypothetical" interventions at age six, i.e. assigning the entire population to one of the trajectory groups that experienced adversity versus the LA trajectory group. The average treatment effects (ATEs) were largest when assigning the population to high LFD compared to LA , with the probability of both diagnoses and convictions being two to three times lower for males and females.

The second largest ATEs were observed when assigning the population to persistent MD compared to LA, with the probability of diagnoses and convictions being two times lower for males and between two to three times lower for women. The ATES were smallest when assigning the population to early MD compared to LA, with the probability of both outcomes being around two times lower for males and females. These results can be interpreted causally under the assumptions of no interference, counterfactual consistency, conditional exchangeability, and correct model specification. I will discuss these assumptions in more detail below.

My results indicate, in line with previous research, that early life adversity is common in high-income countries (Gilbert et al., 2009). My findings suggest that experiencing any adversity, as measured using indicators recorded in administrative registers, increases the risk of disruptive behaviour and crime. I was also able to show that indicators of early life adversity often cluster together, with loss or threat of loss of a family member and negative family dynamics frequently co-occurring in the high LFD group (Briggs et al., 2021; de Vries et al., 2022). Furthermore, individuals who are exposed to these more "severe" adversities show a particularly high risk of experiencing poor psychosocial outcomes.

Similar to previous research on DANLIFE (Bengtsson et al., 2020; El-Khoury et al., 2021; Rod et al., 2020) and other cohort studies (Adjei et al., 2022), some individuals experience higher levels of adversity related to poverty, with two trajectory groups (early MD and persistent MD) showing elevated levels in the material deprivation dimension only. I also found differences between those who were exposed to material deprivation in the first few years of life (early MD) and those exposed throughout early childhood (persistent MD). Individuals who persistently experienced material deprivation showed much poorer outcomes, particularly convictions of sexual and

violent crime, than individuals who only experienced it for the first few years of childhood. This may mean that poverty has a greater impact when experienced later in childhood due to sensitive periods (Bornstein, 1989). Alternatively, these results could represent recency effects, whereby the effect of the exposure on the outcome decreases over time. It could also be that material deprivation has a cumulative effect over childhood and/or that individuals who experience material deprivation from birth to six years continue to experience it after the age of 6. Therefore, our results sustained exposure to material deprivation. Future research should investigate these mechanisms of action.

As well as reporting the potential impact of hypothetical interventions, these findings may be useful for current interventions for DBDs. For instance, the results suggest that certain adversities, such as indicators for negative family dynamics and loss or threat of loss in the family, tend to cluster together. It might be useful to consider these clusters of adversities when intervening in families and acknowledge the added complexity experienced when treating multiple adversities. Furthermore, material deprivation was common in this population. Therefore, a potential intervention which focussed on this dimension (e.g. cash transfers) might have a large impact by increasing financial security, which is reported to have knock-on effects on parental mental health and child outcomes (Akee et al., 2010; Costello et al., 2003, 2010, 2010; McGuire et al., 2022; Thomson et al., 2022; Zimmerman et al., 2021).

Regarding the generalisability of findings, the current study used an entire unselected population, representative of those born in Denmark. However, it did not include individuals who resided in Denmark over the study period but were born elsewhere, i.e. individuals who emigrated to Denmark. Furthermore, it should be noted that rates of more “severe” adversities (e.g. death of a parent or parental drug and alcohol

abuse) are likely to be lower in Denmark compared to other countries, for example, countries with less developed health care systems. Denmark has one of the highest median incomes in the world (national income per capita in 2021 in Denmark = $58,796; European Union average = $31,458; World Bank, 2024), and parental benefits are generous, e.g. an average of 24 weeks of parental leave. Therefore, the current findings may not be transportable to countries with less social security, including high-income countries like the United Kingdom. The population analysed is also highly homogeneous, so the results may not translate to countries with more culturally diverse populations as other factors may influence patterns of adversity, mental health diagnoses, and conviction rates, e.g. racial inequalities in access to resources and healthcare.

6.4.1 Strengths and limitations

The current study has several key strengths. Firstly, using administrative data meant that I could objectively measure adversity. I was also able to include many adversities that are either difficult to assess reliably through questionnaires (e.g. parental alcohol and drug abuse) or have a low prevalence in the population (e.g. death of a parent or sibling). Additionally, as the dataset included everyone born in Denmark since 1980, and the linkage is such that there was extremely limited missing data, the results are highly representative of the target population and, therefore, reliable. Finally, the study captured data from the first six years of life and immediate outcomes, which allowed me to establish the temporal ordering of variables and include a long-term follow-up period for the outcomes.

The current study has some limitations that should be noted. In terms of the analyses, to fit the GBTM, I transformed counts of adversities into binary yearly indicators of each dimension. On the one hand, this meant some information on the severity of

early life adversity was lost. However, using GBTM enabled me to capture the high dimensional structure of adversity over time. Furthermore, GBTM assumes that all individuals assigned to a group follow the same trajectory. Although I evaluated the model fit indices, there is still some uncertainty and risk of misclassification while assigning each child to a latent class based on their posterior probabilities (Herle et al., 2020). This misclassification could potentially bias the survival analysis used to estimate the ATEs (i.e. *counterfactual consistency*). GBTM also assumes no residual correlation among the indicators of early adversities (i.e. *conditional exchangeability*). The dimensions of adversity were constructed based on an expert-derived framework; therefore, I did not examine correlations and direct effects between specific early adversities. Future research could examine the relative impact of each adversity on disruptive behaviour while controlling for the other adversities, using target trial emulation (Hernán & Robins, 2016). Finally, all analyses also assumed that observations were independent (i.e. *no interference*), which may not be true given that siblings and parents are included in this population. It might be interesting to investigate whether impacts results by clustering the analyses by families or using a subsample of individuals.

Although using administrative data enabled us to follow over 1.9 million individuals throughout adolescence and adulthood, registry data also have limitations. For instance, some information on key adversities, such as child physical and sexual abuse, was unavailable in DANLIFE. However, these adversities may be indirectly captured through other measured adversities, such as foster care. Moreover, the number of certain adversities (e.g. parental drug and alcohol use) is likely to be higher than those recorded in administrative datasets. Furthermore, the data also excludes individuals who exhibit subclinical levels of disruptive behaviour, exhibit clinical levels of disruptive behaviour but never receive a diagnosis, commit a sexual or violent crime but are never caught by police or who are caught by police but are never

convicted. Therefore, the prevalence of these outcomes in the current study (0.3 % for diagnoses and 1.8% for convictions) are much lower than the actual rates of disruptive behaviour in the population. Consequently, the current study may underestimate the impact of early life adversity on disruptive behaviour.

6.4.2 Conclusions

I found that individuals born in Denmark since the 1980s cluster into four latent groups of early life adversities: low adversity, high levels of loss or threat of loss and negative family dynamics, early material deprivation and persistent material deprivation. Compared to individuals in the low adversity group, individuals in the other three trajectory groups who experienced early life adversity had a greater risk of being diagnosed with disruptive behaviour disorders or committing a sexual or violent crime by the age of 25 years. A hypothetical intervention that assigned individuals to the lowest adversity group, if such an intervention existed, could lead to a three-fold reduction in the probability of diagnoses and convictions before the age of 25 years. The current findings demonstrate the potential benefit of interventions that reduce exposure to adverse childhood experiences on levels of disruptive behaviour, including diagnoses of CD and DPD and convictions of violent and sexual crimes.

Key points

1. In this Chapter, I used administrative data on 1,900,369 children from an unselected population-based cohort to investigate the impact of early life adversity on diagnoses of conduct disorder (CD) and dissocial personality disorder (DPD) and convictions of sexual and violent crimes.

2. I grouped individuals based on whether they had experienced indicators of adversity on three dimensions (material deprivation, loss or threat of loss within the family, negative family dynamics) from birth to age six years.

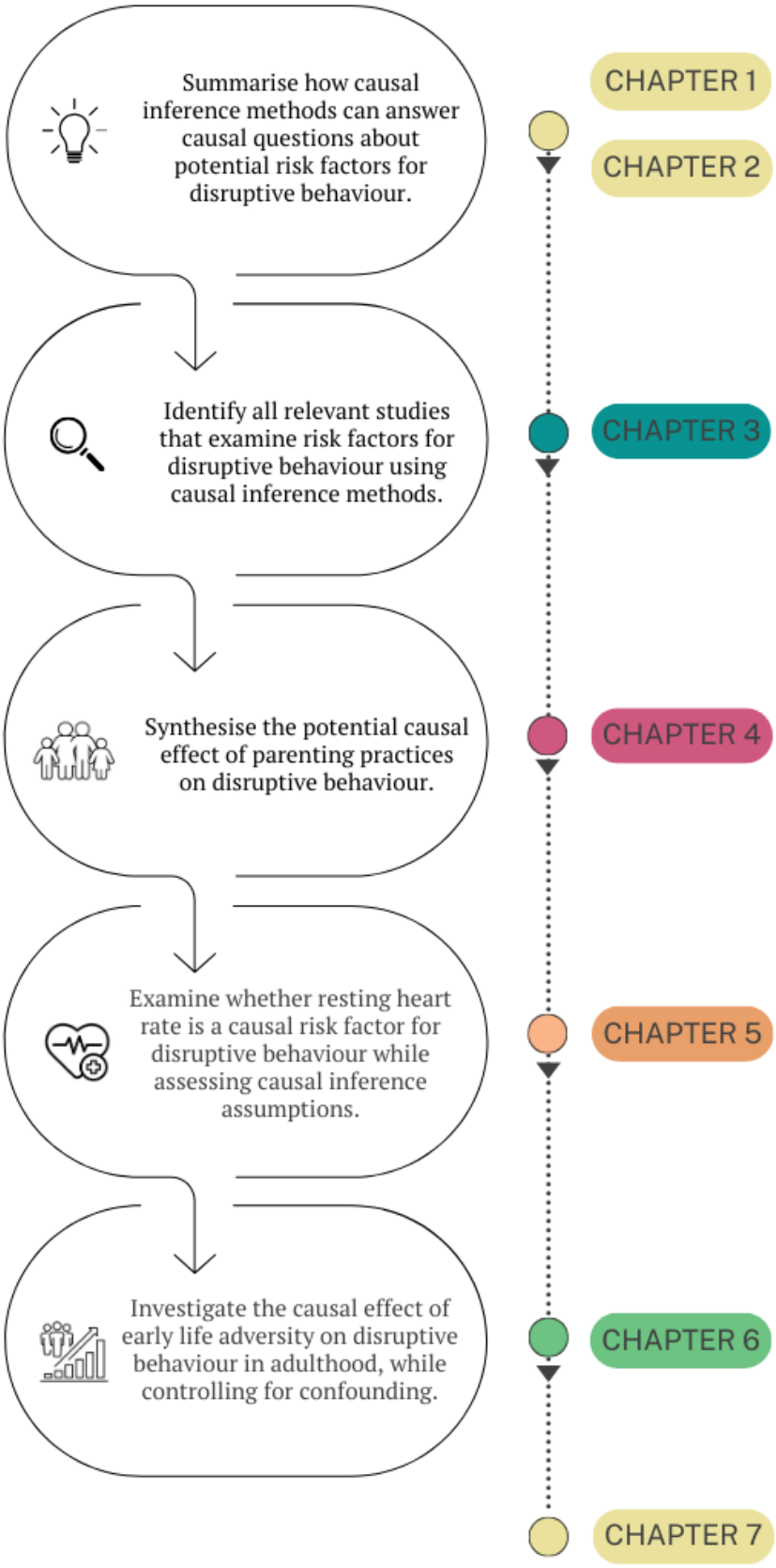
3. I identified four latent trajectory groups: low adversity (*n* = 1,392,151; 68.9% [of the population]), early material deprivation (*n* = 226,301; 11.2%), persistent material deprivation (*n* = 288,937; 14.3%), high loss and negative family dynamics (*n* = 113,151; 5.6%).

4. The rates of diagnoses and convictions were higher for individuals in groups that experienced early life adversity than those in the low adversity group.

5. I estimated the average treatment effects (ATEs) of a hypothetical intervention that assigned individuals to low adversity and predicted that, if such an intervention were to exist, it could lead to a two- to three-fold decrease in the probability of diagnoses and convictions by the age of 25 years.

6. These findings highlight the potential benefits of interventions that reduce exposure to early life adversity on levels of disruptive behaviour, including diagnoses of CD and DPD and convictions of violent and sexual crimes.

THESIS STRUCTURE



7 DISCUSSION

Chapter overview

In this thesis, I have summarised the main causal inference methods for assessing the causal effect of risk factors on DBDs. Using different causal inference methods, I have triangulated evidence for three selected putative risk factors for DBDs (parenting practices, resting heart rate and early life adversity). I quantitatively summarised the causal effect of parenting practices on DBDs in a meta-analysis of studies using causal inference methods. I examined whether the often-reported association between resting heart rate and antisocial behaviour was causal using Mendelian randomisation analyses. I estimated the impact of a hypothetical intervention that reduced early life adversity on DBDs using g-computation. In this Chapter, I will summarise the key findings presented in this thesis, review how the empirical Chapters sit together, build upon the existing literature and lay the foundations for future research. I will also outline this thesis's strengths and limitations and detail important potential next steps.

7.1 Summary of key findings

In this thesis, I set out to triangulate evidence of potential causal risk factors for disruptive behaviour disorders (DBDs) by evaluating the strength and reliability of the estimates reported in existing research, using large representative datasets, and implementing novel causal inference methods.

To identify gaps in the existing literature, I conducted a systematic review of research using causal inference methods to investigate the aetiology of diagnoses and symptoms associated with DBDs. In the study protocol (Chapter 2, also Karwatowska et al., 2020), I defined key terms used in the causal inference literature and categorised causal inference methods into two categories, specifically those that aim to estimate causal effects by *(a)* relying on an instrument (e.g. regression discontinuity, Mendelian randomisation (MR), difference-in-difference approaches) or by *(b)* confounder-control (e.g. extensions to regression-based methods, propensity score matching).

I identified 167 studies published between 1980 and 2021 that examined 23 putative risk factors for disruptive behaviour in 934,876 individuals from 73 distinct cohorts across 18 countries using 13 causal inference methods (Chapter 3). These findings demonstrate the breadth of research in this area. Combining evidence from multiple sources can aid triangulation by identifying gaps in the causal inference of specific risk factors and highlighting the scarcity of certain methods. The majority of research has been focussed on three risk factors (parenting practices [k [number of studies] = 39], prenatal exposure to toxins [k = 26] and parental internalising symptoms [k = 17]) and three causal inference methods (propensity score matching analyses [k = 35], discordant sibling study design [k = 32] and adoption study [k = 29]). However, most of the 23 risk factors had been examined by fewer than ten studies, indicating a clear need for more research using causal inference methods. Furthermore, most methods used were family-based, such as sibling and adoption studies, and propensity score matching analyses. It is essential to diversify the causal inference methods used to examine risk factors and use underutilised methods such as MR and g-methods.

The risk factors that have been examined using a variety of causal inference methods could be considered for quantitative syntheses via meta-analyses. To triangulate

evidence on a risk factor targeted in preventative interventions for DBDs (i.e. Parent Management Training; PMT), I conducted the first meta-analysis of studies that examined parenting practices using causal inference methods (Chapter 4). I updated and restricted the searches detailed in Chapter 3 and identified 41 studies on parenting practices that included data from 27 distinct cohorts with 36,661 individuals. The findings suggested that negative parenting practices may have a causal effect on offspring DBD symptoms. The estimate varied depending on the type of causal inference method used, the study quality and whether the same informant rated the exposure and outcome. Therefore, only the highest-quality studies were included in my final pooled meta-analytic estimate ([pooled Pearson's] = 0.104; *95% CI* = 0.053, 0.154; *n* = 16,101). The results indicated that a unit increase on a standardised measure for negative parenting practices (such as the Iowa Family Interaction Rating Scales; Williamson et al., 2011) was associated with an increase of 0.102 SD on a standardised measure for symptoms of DBDs (such as the Strength and Difficulties Questionnaire; Mieloo et al., 2012). To further quantify this causal effect, I estimated the population attributable impact of decreasing negative parenting practices by 0.4SD (the magnitude of change reported in universal prevention programmes, Jeong et al., 2021). I estimated that if the pooled meta-analytic estimate reflects the true causal effect of negative parenting practices, a hypothetical intervention applied globally could lead to a 0.11% reduction in the prevalence of clinically relevant DBD symptoms worldwide, the equivalent of 3,614,337 school-aged children no longer exhibiting clinical levels of DBD symptoms (Polanczyk et al., 2015).

In Chapter 5, I conducted a two-sample Mendelian randomisation (MR) analysis using genetic variants associated with resting heart rate (RHR) as instrumental variables (IVs) to explore whether RHR is a *"possible causal risk factor"* (Portnoy & Farrington, 2015, p. 42) for antisocial behaviour (ASB). I reported no evidence of causal effects for

RHR on ASB in the MR analyses (N_{SNPs} [number of SNPs] = 278; B_{IVW} [inverse variance weighted estimate] = –0.0004; *95% CI* = -0.004, 0.004). I also carried out further sensitivity analyses using linkage disequilibrium score regression (LDSC), which suggested no evidence of significant genetic correlations (r_{GE}) between RHR and ASB (r_{GE} = 0.057, p = 0.169). Although the findings from this study do not rule out other associational pathways between RHR and ASB, they do not support the theory that the relationship between RHR and ASB is directly causal.

In Chapter 6, I analysed a whole population-based cohort of over 1.9 million individuals to study the effect of early life adversity on DBDs. Using data on three dimensions of adversity from birth to age six, I identified the presence of four trajectory groups: low adversity (LA), high loss and negative family dynamics (high LFD), early material deprivation (early MD), and persistent material deprivation (persistent MD). The findings revealed that individuals exposed to early life adversity had a higher risk of being diagnosed with either conduct disorder (CD) or dissocial personality disorder (DPD) or being convicted of a sexual or violent crime. Individuals who experienced the highest levels of adversity (high LFD) showed the highest rate of diagnoses and convictions. Compared to the LA group, males in the high LDF group exhibited 2.7 times the rate of diagnoses ([adjusted hazard ratio] *a-HR* = 2.65; *95% CI* = 2.39, 2.93) and 2.4 times the rate of convictions (*a-HR* = 2.39; *95% CI* = 2.27, 2.50). Females showed 3.2 times the rate of diagnoses (*a-HR* = 3.17; *95% CI* = 2.67, 3.77) and 3.1 times the rate of convictions (*a-HR* = 3.12; *95% CI* = 2.77, 3.52) compared to the LA group. I estimated the potential outcomes of assigning the population to the high LFD group versus the LA group. The average treatment effect (ATE) suggested that a hypothetical intervention that assigned individuals to the LA group would lower the probability of diagnoses and convictions by 2.7 and 2.4 times for males and 3.2 and 3.1 times for females, respectively.

I have extracted three main themes from the results of this thesis. I will discuss the following themes in the sections below: *(i)* possible aetiological mechanisms underlying DBDs, *(ii)* the impact of hypothetical interventions for DBDs, and *(iii)* triangulation using different causal inference methods.

7.1.1 Possible aetiological mechanisms

The current findings of causal (and non-causal) effects indicate potential mechanisms of action underlying the development of DBDs. Regarding evidence of non-causal effects, the results from Chapter 5 suggest no direct causal effect of RHR on ASB. These findings were supported by additional analyses, including a multivariable MR analysis and LDSC, which examined the effect of alternative risk factors (resting heart rate variability; HRV) and the presence of potential genetic confounding between RHR and ASB. The additional analyses did not suggest alternative risk factors associated with heart rate were causally related to ASB nor any genetic overlap between any measure of heart rate and ASB. Two previous studies using genetically informed methods also found no effect of RHR on DBDs in childhood (Baker et al., 2009) and adulthood (Kendler et al., 2021), and a previous study using LDSC also found no genetic correlation between RHR and childhood aggression (Ip et al., 2021). These results suggest that RHR does not directly contribute to ASB.

Therefore, previous findings of an association between RHR and ASB must be driven by other mechanisms. For example, one possibility is that RHR could be associated with other differences in autonomic nervous system (ANS) activity, which may contribute to ASB. This could result in a non-causal association between RHR and ASB (i.e. ANS → RHR and ANS → ASB). Other potential common causes of low RHR and high ASB could be sensation-seeking behaviour (Hammerton et al., 2018; Portnoy & Farrington, 2015; Sijtsema et al., 2010). Although the results from the MR analyses in Chapter 5 cannot clarify these alternative mechanisms, they indicate that RHR is

not a risk factor for DBDs. However, RHR may still be used as a non-causal associational marker of DBDs, which I will discuss in more detail below in Section 7.2.1.

Regarding evidence of potential causal effects, the meta-analytic results presented in Chapter 4 suggest that negative parenting practices *do* have a small causal effect on offspring DBDs. This finding is consistent with evidence from randomised control trials (RCTs) on the main intervention for DBDs, which indicate that PMT has a moderate effect on child disruptive behaviour (Cohen's d = -0.21 to -0.69; (Leijten et al., 2019, 2022). My findings also indicated that the effect of parenting was similar for mothers and fathers, suggesting similar or shared mechanisms for maternal and paternal parenting. Additionally, the effect of parenting was consistent across males and females, indicating no sex differences in the effect of negative parenting. This implies that the higher prevalence of DBDs in males cannot be explained by differences in the effects of parenting practices (Lysenko et al., 2013; Pinquart, 2017). Finally, the effect of parenting practices was stable over the range of ages examined in the studies (0 – 37 years), suggesting that there may not be key (or "sensitive") periods during which a child is more susceptible to negative parenting practices (Gardner et al., 2019). However, as most of the studies only included "normative" negative parenting practices, this finding does not preclude the presence of sensitive periods for extreme forms of negative parenting, such as maltreatment and neglect.

The finding that the effect of parenting practices varies based on the causal inference method used might also indicate the presence of different genetic and environmental confounders as causal inference methods control for different types of confounding. For example, adoption studies and discordant sibling studies reported the largest effects. Adoption studies control for genetic confounding but may not always control for prenatal differences, such as exposure to toxins or symptoms of psychopathology,

which are higher in women whose children are adopted (Gaysina et al., 2013). On the other hand, sibling studies do not control for non-shared environmental factors, such as peer relationships or other factors that might influence parental parenting practices (Frisell, 2020). The meta-analysis in Chapter 4 is an example of how evidence from different causal inference methods can be triangulated to examine potential genetic and environmental factors that require future research.

In Chapter 6, I replicated previous findings that different types of early life adversities often co-occur in the same individuals (Gilbert et al., 2009). Although I did not directly investigate the mechanisms by which these adversities cluster, recent research has attempted to do this using network analyses (de Vries et al., 2022). I also found evidence that individuals who are exposed to more "severe" adversities (e.g. death of a parent, parental substance abuse) are at a particularly high risk of experiencing poor psychosocial outcomes. Therefore, different types of adversities may carry different risks and influence specific outcomes and research using multiple outcomes could investigate this further (Adjei et al., 2022; Baldwin et al., 2023). The findings that early life adversity is associated with DBDs using prospectively recorded objective measures also suggest that this relationship is not entirely driven by biases in recall or selection.

I also found differences between the two groups that experienced MD. Individuals who experienced early MD had much lower rates of diagnoses and convictions compared to individuals who were exposed to persistent MD. These results may reflect a "recency" effect, whereby the impact of an exposure is larger for proximal (or more recent) rather than distal events (Shanahan et al., 2011). Alternatively, these results may suggest a "cumulative" effect of MD, whereby there is a dose-response relationship with the number of years exposed to early life adversity (Evans et al., 2013). Finally, individuals in the persistent MD group may continue to experience MD

after the age of six. Although the current findings cannot disentangle these potential mechanisms, future research using DANLIFE could look at the direct causal effect of each indicator of adversity while controlling for the other correlated adversities using the target trial emulation (TTE) framework (Hernán & Robins, 2016; Matthews et al., 2022; discussed further in Section 7.4.1).

7.1.2 Hypothetical interventions

Although this thesis does not directly test the efficacy of interventions for DBDs, where I uncovered evidence of causal effects, I have attempted to estimate the impact of hypothetical interventions. In Chapter 4, I attempted to quantify the potential causal effect of parenting practices on disruptive behaviour by estimating the "population attributable impact" of negative parenting on the prevalence of DBDs. To estimate this figure, I relied on previous estimates of the global prevalence of DBDs (Polanczyk et al., 2015) and the change to negative parenting practices following universal prevention programmes (Jeong et al., 2021). Using these estimates, I predicted that an intervention reducing negative parenting behaviours could lead to a 0.11% reduction in the prevalence of clinically relevant DBD symptoms worldwide (n = 3,614,337). This suggests that even a small causal effect (r = 0.104) can impact the development of DBDs and demonstrates that even a small shift in the population mean can result in a clinically significant drop in cases. Due to the long-term adverse consequences of DBDs, preventing even a small fraction of the population from developing these symptoms is expected to have large and positive downstream consequences (Burt et al., 2018).

However, it should be noted that strong assumptions are invoked when calculating the population attributable impact. As well as the causal inference assumptions discussed in more detail in Section 7.3.5, the calculation of the population attributable impact also relies on the correct calculation of the estimates used (i.e.

estimates of the prevalence of DBDs, the impact of interventions on negative parenting practices and the effect of negative parenting on DBDs estimated from my meta-analysis). Furthermore, it is also assumed that the populations used to calculate these estimates are comparable, which may not be justifiable. Therefore, the population attributable impact should not be overinterpreted. Instead, it should represent an example of how researchers can convey the potential impact of their results, making them more accessible to readers (Funder & Ozer, 2019).

In Chapter 6, I estimated the impact of hypothetical interventions on early life adversity using the g-computation formula (Gerds et al., 2023; Robins, 1986). I predicted that an intervention which meant that all individuals in the population were exposed to LA (compared to high LFD) would reduce the probability of diagnoses and convictions by two times for males and over two times for females. This suggests the possible benefit of intervening on early life adversity. G-computation allows researchers to estimate the marginal, as opposed to the conditional, effect of an exposure and can be interpreted as causal in situations where certain assumptions hold. However, as I will discuss in more detail in Section 7.3.5, it is difficult to justify these assumptions in situations where there is uncontrolled confounding (no *exchangeability*), when the intervention is poorly defined (no *consistency*), or when the probability being exposed is either 0 or 1, conditional on the covariates (no *positivity*). Furthermore, although g-computation estimates population effects, they are only representative of the target population. Nevertheless, g-computation represents a robust but underutilised method for estimating marginal population causal effects in epidemiology. It is a useful tool for researchers to examine the potential impact of hypothetical interventions that target specific risk factors which are difficult or impossible to investigate using RCTs.

7.1.3 Triangulation using different causal inference methods

This thesis contains several examples of how triangulation of existing evidence can advance our knowledge of the mechanisms of DBDs. In order to draw causal conclusions, triangulation of evidence is needed using methods that invoke different but complementary assumptions (De Stavola et al., 2022; Lawlor et al., 2017; Munafò & Smith, 2018). In my review, I identified gaps in the literature where further triangulation is needed using more diverse causal inference methods to examine potential risk factors for DBDs. I identified that, although there had been a handful of studies using genetically informed methods to examine the effect of RHR on ASB, no previous research had examined RHR using MR analyses. The null findings illustrate that novel causal inference methods can be used to disprove, as well as prove, theories in aetiological epidemiology. Furthermore, Chapter 4 showed that it is possible to quantitatively triangulate evidence by synthesising estimates from causal inference methods in meta-analyses. Finally, in Chapter 6, I illustrated the utility of triangulating using different methods *and* different types of data to address biases inherent to specific risk factors, i.e. recall bias when retrospectively reporting early life adversity.

7.2 Translational implications

The results from this thesis have potential clinical, public health and methodological implications.

7.2.1 Clinical implications

Although this thesis does not directly explore the effectiveness of interventions, the results could have clinical implications, particularly for universal prevention programmes in the general population. The finding of a small harmful causal effect of negative parenting practices should encourage clinicians to attempt to reduce

these behaviours. The effect of parenting practices was much lower than those previously reported in RCTs of PMT (Mingebach et al., 2018), which may suggest that PMT influences DBDs not only through parenting practices but also through influencing other risk factors, such as parental relationship quality and psychopathology (Jeong et al., 2021; Weber et al., 2019). The studies in my meta-analysis examined specific parenting practices, and most attempted to control for other related variables. Even though fathers are underrepresented in parenting interventions, our results indicate that the effect of parenting practices is similar for mothers and fathers. Therefore, paternal involvement in interventions may have a beneficial effect in preventing offspring DBD symptoms (Lundahl et al., 2008; Panter-Brick et al., 2014).

The results of Chapter 6 added to a large body of evidence indicating that early life adversities cluster together. Previous researchers have suggested that it might be useful to consider these clusters of adversities when intervening in families. For example, when clinicians work with individuals exposed to one adversity, knowing to check for frequently co-occurring adversities. Furthermore, the clustering of adversities may, in part, reflect a genetic vulnerability that correlates with environmental risks (i.e. r_{GE}). It is important to acknowledge that r_{GE} adds to the complexity of working with families that are experiencing multiple adversities and where parents and children may both have disruptive behaviour/temperament.

Finally, it is important to note that even if causal inference methods show that a factor is not causally related to DBDs, they may still have clinical utility. For example, RHR may be associated with ASB through non-causal pathways. Nonetheless, if RHR is robustly associated with ASB, then it could be used as an indicator of ASB, such as in risk assessments and a measure of the change after behavioural interventions (de Looff et al., 2022; Portnoy & Farrington, 2015). Even though causal inference

methods on observational data are useful when RCTs are impractical or unethical, as is the case for many putative risk factors for DBDs (e.g. early life adversity), they do not replace the need for RCTs to assess the implementation and efficacy of intervention for DBDs.

7.2.2 Public health implications

Although RCTs are vital to test pragmatic treatment options, causal inference methods applied to observational data are important in informing public health policies. For instance, RCTs on smoking were not ethical, and therefore, smoking reduction policies were based largely on the convergence of observational studies in humans and experimental animal studies (Pingault et al., 2018).

The results of this thesis suggest a potential benefit of universal interventions that reduce negative parenting practices (population attributable impact) and interventions that reduce exposure to early life adversity (g-computation formula). Regarding interventions on early life adversity, although it may be difficult to intervene on certain adversities (e.g. loss or threat of loss in the family), I predicted the potential outcomes following interventions on MD (i.e. poverty and parental long-term employment), which also indicated a reduction in the probability of diagnoses of CD and DPD and convictions of sexual and violent crimes. This adds to a body of research which suggests that cash transfers may increase mental health and well-being (McGuire et al., 2022; Thomson et al., 2022; Zimmerman et al., 2021) and decrease levels of disruptive behaviour and crime (Akee et al., 2010; Costello et al., 2003).

When discussing possible implications to public health, it is also important to acknowledge that policy changes should not be made by a single study (Bann et al.,

2024) and that there are many potential barriers to research impact. For example, research on mental health, and particularly child mental health, is chronically underfunded (MQ Transforming Mental Health., 2021). This is reflective of current public opinion and the political landscape. Therefore, even if better and more robust research unequivocally showed a link between universal parenting programmes (or increases in financial aid) and better child mental health outcomes, would current governments "do something, do more, do better" (Marmot, 2020)?

7.2.3 Methodological implications

In terms of this thesis's methodological implications, each Chapter showcases different ways that causal inference methods can be used to examine the aetiology of DBDs. Chapters 1 and 2 summarise the key biases in this area and describe the causal inference methods that can be used to address these biases. The results from the systematic review in Chapter 3 identified key gaps in the causal inference literature, which can inform future research. For example, few studies have used inverse probability of treatment weighting (IPTW), MR and natural experiments. The meta-analytic results reported in Chapter 4 suggested the presence of shared method variance, the bias induced when the same informant reports the exposure and outcome measures, as the pooled estimate was much lower when different reporters were used (Francis et al., 2023; Podsakoff et al., 2003; Schoeler et al., 2018), which underlines the importance of using multiple informants and objective measures, e.g. administrative data. Furthermore, the findings that the effect varied by the causal inference method used confirm that although causal inference methods can theoretically minimise certain confounding biases, in practice, no one causal inference method can fully succeed in doing so in isolation (Goetghebeur et al., 2020; Lawlor et al., 2017; Munafò & Smith, 2018). Therefore, researchers need to triangulate in order to identify potential sources of bias.

7.3 Strengths and limitations

This thesis has many strengths, including using large and representative datasets, a diverse set of risk factors and causal inference methods, and extensive sensitivity checks to falsify key causal inference assumptions. However, it is important to discuss its limitations, such as limited ability to detect small effects, heterogeneity in measures used, homogeneity of the study populations, reliance on secondary data, and potential violations of the causal inference assumptions.

7.3.1 Power

I used large pre-existing datasets in each one of my empirical Chapters, including data on 36,661 individuals for the meta-analysis (Chapter 4), 458,835 individuals for the exposure and 85,359 individuals for the outcome in the MR analyses (Chapter 5) and 1,900,369 individuals for the g-computation analyses (Chapter 6). However, it may be that these studies were still not sufficiently powered to detect small effects. However, the clinical utility of such small causal effects is uncertain. As sample sizes continue to increase, these analyses could be updated.

7.3.2 Heterogeneity

Another potential limitation of the current thesis is the use of heterogeneous measures, which reduces power, biases the results and makes interpretation difficult. In the MR analyses (Chapter 5), the outcome GWAS combined different measures of ASB collected from multiple cohort studies. Similarly, in Chapter 4, I combined estimates from a wide range of phenotypes for the exposure (parenting practices) and the outcome (DBD symptoms) in the meta-analysis. However, subgroup analyses showed that the meta-analytic effect was consistent for different outcomes. I also synthesised estimates from many different causal inference methods. Causal

inference methods estimate different estimands and have different target populations. This can make synthesising the estimates from different methods challenging and potentially problematic (Becker et al., 2017). However, researchers can make sound causal inferences after carefully considering the estimands and target populations and/or using advanced meta-analytic methods (D. Jackson et al., 2024).

7.3.3 Generalisability

Although there are exceptions, historically, there is a striking lack of diversity in epidemiological research. Indeed, there is a lack of demographic, geographic and ancestral diversity in cohorts (e.g. biases in recruitment and differential attrition), genetic datasets (e.g. use of homogenous and exclusively European samples) and available data (e.g. focus on high-income countries; HIC). The datasets used in this thesis were large, including two datasets combined from multiple sources and one including an entire population (Denmark). However, the ancestry of individuals within those datasets was still fairly homogeneous, i.e. the majority were from White European ancestry. Furthermore, the results of my systematic review (Chapter 3) highlight the imbalance of research in HIC. Even though the 167 studies included data from 18 countries, all are HIC. Furthermore, in my meta-analysis (Chapter 4), most studies either did not report information on ancestry or included samples that were primarily from White ancestry.

As this thesis relies on secondary data, the lack of diversity is also apparent in the MR and g-computation analyses. For example, in Chapter 5, the GWAS for the exposure (RHR) and the outcome (ASB) solely included data from individuals of European descent. Furthermore, 94.4% of the DANLIFE cohort had parents of European descent (Chapter 6), and I could not include individuals born outside of Denmark who later emigrated to Denmark. The lack of diversity in epidemiology makes generalisability and applicability difficult and further exacerbates inequality in research.

7.3.4 Secondary data analysis

In addition to the limitations already mentioned, the reliance on secondary data also limited the scope of the risk factors that were examined. For example, the availability of GWAS with SNP-exposure associations restricted the potential risk factors that could be examined using MR. Furthermore, using administrative datasets meant that some key risk factors, such as child physical and sexual abuse, could not be considered as they were not recorded in the dataset. Administrative data is further restricted to official records and there is a large discrepancy between the actual rates of adversity and those recorded in the registries (Gilbert et al., 2009). The same is true for diagnoses and convictions, with actual rates of clinically relevant symptoms of DBDs higher than those captured in official records. As such, it is crucial to triangulate evidence using multiple information sources as well as different methods (De Stavola et al., 2022; Lawlor et al., 2017; Munafò & Smith, 2018).

7.3.5 Causal inference assumptions

The most important limitation of this thesis is the potential violations of its underlying assumptions. All causal inference is based on strong assumptions, such as *no interference, counterfactual consistency, conditional exchangeability*, and *positivity* (Hernán & Robins, 2020). Further assumptions of relevance, exchangeability and exclusion restriction are invoked when using an IV. However, these assumptions are often impossible to verify and must be falsified instead (Bärnighausen, Oldenburg, et al., 2017). If these assumptions are not met, any findings will be biased and, by definition, not causal.

Potential violations of these assumptions in traditional two-sample MR analyses are increasingly being recognised and discussed in the literature (Sanderson et al., 2022 for a review). There are many situations when they might be violated, including in

situations where SNPs are weak instruments for the exposure (violation of *relevance;* in situations where there is population stratification (i.e. when ancestry influences genetic variants and the phenotype), dynastic effects (i.e. when genetic variants of the parent influence the phenotype of the child) or assortative mating (i.e. when individuals select partners who are similar to them; violations of *exchangeability*; Brumpton et al., 2020), and in situations where there is pleiotropy (i.e. when genetic variants have effects on multiple phenotypes) or linkage disequilibrium (i.e. non-random associations between genetic variants; violations of *exclusion restriction;* Hemani et al., 2018). Further, more nuanced assumptions include that any changes in genetic variation are equivalent in their effects to changes in the exposure through environmental or pharmaceutical manipulation (gene-environment equivalence) and that the exposure and outcome GWAS have the same underlying populations (Sanderson, Glymour, et al., 2022). Although I ran numerous sensitivity analyses, I could not falsify all of these assumptions. However, MR methods that are more robust to these biases are constantly being developed (Hwang et al., 2021), creating exciting opportunities for future research in this area.

The assumptions underlying the g-computation formula could also be violated in this thesis. For example, the no interference assumption (i.e. that one individual's exposure level does not influence another's outcome) may be violated when samples include both parents and children. It is plausible that the exposure level of a parent (e.g. high adversity) could affect the outcome of their children (e.g. disruptive behaviour). As DANLIFE includes data on families, analyses could be run on subgroups (i.e. without family members) and/or with a clustering variable for families. The assumption of conditional exchangeability (i.e. exposed and unexposed individuals are exchangeable, conditional on covariates), the equivalent of no (unmeasured) confounding, is also difficult to justify when dealing with complex relationships between exposure and outcomes. Along with using more robust

methods and carefully considering the data structures, researchers should include directed acyclic graphs (DAGs) to visualise the underlying assumptions of their analyses and increase transparency.

The counterfactual consistency assumption (i.e. that the observed outcome is the same as the potential outcome that would have occurred if the exposure level were set by the researcher to the same value as the observed exposure level) is particularly pertinent when using constructed psychosocial measures, such as questionnaires and latent variables (VanderWeele, 2022). Using the example of the latent trajectory groups in Chapter 6, the pathways that led individuals to be exposed to the indicators of early life adversity will vary hugely, allowing for many different versions of interventions (VanderWeele, 2009). In other words, there are multiple possible interventions that could lead to individuals being assigned to the LA group. Researchers should be aware of possible violations of the consistency assumption when designing their study and interpreting the results. TTE, described in Section 7.4.1, provides a useful formal framework that encourages researchers to consider common biases in observational data (Hernán & Robins, 2016).

Each causal inference method has specific uses and limitations and should be chosen depending on the causal question of interest. Although no method is perfect, researchers can be more confident about the relationships between risk factors and DBDs by comparing the results from causal inference methods with different and, hopefully, complimentary sources of bias.

7.4 Next steps and future research directions

The work outlined in this thesis generates many questions that could be answered in further research by extending the methods used, including other risk factors and/or comorbid outcomes, and using more diverse samples and populations.

7.4.1 Extensions to the current work

In terms of extending the work outlined in this thesis, Chapter 3 identified the potential risk factors that have been examined using causal inference methods and could be triangulated via meta-analysis, similar to the one outlined in Chapter 4. Two previous reviews have investigated the causal effect of prenatal smoking, alcohol and caffeine exposure (Haan et al., 2022) and childhood maltreatment (Baldwin et al., 2023) on DBDs. The other risk factors identified in my systematic review include prenatal antidepressant exposure, parental internalising symptoms (e.g. depression, stress and anxiety), family dynamics (e.g. parental separation, marital conflict), education (e.g. childcare) and parental substance (ab)use (e.g. drugs and alcohol). The systematic review also showed that very few studies have used MR analyses to examine risk factors for DBDs. Therefore, the MR analyses could be extended using other exposures (e.g. testosterone) and/or more advanced MR methods (e.g. within-family MR analyses; Hwang et al., 2021). Finally, the DANLIFE cohort provides a rich and representative dataset. Future research using DANLIFE could use the TTE framework to examine each specific indicator of early life adversity. TTE emulates the framework used in RCTs but with observational data (Hernán & Robins, 2016). It provides a structured approach to designing the "ideal" target trial (e.g. eligibility criteria, treatment strategies, assignment procedures, follow-up period, outcome, causal contrast of interest, analysis plan). Consequently, it helps to avoid self-inflicted biases, such as the prevalent users and immortal time biases (Hernán et al., 2016). A possible target trial could examine the effect of MD on DBDs using g-computation in children whose parents were in poverty for the year before their birth,

controlling for other correlated indicators of adversity. The potential outcome of staying in poverty could be compared to the outcome of no longer being in poverty at different ages. Alternatively, instead of a binary measure of poverty, potential outcomes could be estimates for different family income levels (e.g. increases of 500 Danish krone per month).

7.4.2 Other risk factors

Beyond the risk factors considered in this thesis (parenting practices, RHR and early life adversity), it will be important for future research to consider risk factors for which there is less causal evidence. For example, some potential candidates, which I identified as having been examined in fewer than ten causal inference studies (Chapter 3), included neighbourhood characteristics, parental externalising symptoms, socioeconomic position, parental education and employment, birth weight, breastfeeding, own substance use, parental ACEs, peer relationships, maternal age at birth, adolescent childbearing, gaming, obstetric complications, residential mobility, TV viewing and temperament. This thesis did not consider extra-familial factors, such as peer problems and neighbourhood characteristics. Data on these factors is not always collected in cohort studies (peer problems) or requires access to specific administrative datasets (neighbourhood characteristics). Future research could look at using causal inference methods on these risk factors.

7.4.3 Comorbid outcomes

Increasingly, research suggests that risk factors for DBDs are transdiagnostic, meaning that they confer risk for multiple psychopathologies, such as attention deficit hyperactivity disorder and internalising disorders. This suggests that there is a significant overlap in the aetiology of child mental health. The current thesis exclusively focuses on DBDs. However, it would be interesting to include other types

of psychopathology as outcomes in future analyses. Although including multiple outcomes is limited to the data available, this data is available within administrative datasets (such as DANLIFE) and may identify different mechanisms that influence mental health in children and adolescents.

7.4.4 Diversity and inclusion

Changes are needed to improve recruitment and participation to increase diversity in research. Currently, there is inherent selection bias (and therefore collider bias) in research (Munafò et al., 2018). Researchers should provide information on the ancestry of their study populations so that this information can be included when triangulating evidence. Furthermore, as funders increasingly encourage more diversity (Wellcome, 2021), researchers should include data from LMICs in their grant applications to ensure that evidence and subsequent evidence-based practice include everyone in society, not only the most privileged.

7.5 Conclusions

Based on a systematic review and meta-analysis of the existing causal inference literature and further analyses using two novel causal inference methods, this thesis suggests that parenting practices and early life adversity have a causal effect on disruptive behaviour disorders (DBDs). Although some causal inference methods, such as adoption and discordant sibling study designs, have been frequently used in research on disruptive behaviour, other methods, such as IV analyses and g-methods, are still underutilised. Causal inference methods theoretically provide a quick and efficient way for researchers to identify risk factors for disruptive behaviours, but researchers should be mindful of the underlying causal inference assumptions and limitations of the methods, datasets, and their combinations. Triangulating evidence from various study designs, analytic methods and information sources is essential.

Experimental studies are necessary to test pragmatic treatment options. However, causal inference methods applied to observational data are crucial when RCTs are impractical and unethical, as is so often the case for factors affecting child mental health. Causal inference methods can improve our understanding of risk factors for disruptive behaviour by identifying candidates for targeted interventions and public health initiatives and estimating potential outcomes.

References

Adjei, N. K., Schlüter, D. K., Straatmann, V. S., Melis, G., Fleming, K. M., McGovern, R., Howard, L. M., Kaner, E., Wolfe, I., & Taylor-Robinson, D. C. (2022). Impact of poverty and family adversity on adolescent health: A multi-trajectory analysis using the UK Millennium Cohort Study. *The Lancet Regional Health - Europe*, *13*, 100279. https://doi.org/10.1016/j.lanepe.2021.100279

Akee, R. K. Q., Copeland, W. E., Keeler, G., Angold, A., & Costello, E. J. (2010). Parents' Incomes and Children's Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits. *American Economic Journal: Applied Economics*, *2*(1), 86–115. https://doi.org/10.1257/app.2.1.86

Allen, N. E., Sudlow, C., Peakman, T., & Collins, R. (2014). UK Biobank Data: Come and Get It. *Science Translational Medicine*, *6*(224), 224ed4. https://doi.org/10.1126/scitranslmed.3008601

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: Fifth Edition (DSM-5)* (Vol. 5, Issue 5). American Psychiatric Association Washington, DC.

Anderson, S., & Leventhal, T. (2017). Residential mobility and adolescent achievement and behavior: Understanding timing and extent of mobility. *Journal of Research on Adolescence*, *27*(2), 328–343. https://doi.org/10.1111/jora.12288

Anthony, R. E., Paine, A. L., & Shelton, K. H. (2019). Adverse Childhood Experiences of children adopted from care: The importance of adoptive parental warmth for future child adjustment. *International Journal of Environmental Research and Public Health*, *16*(12), 2212. https://doi.org/10.3390/ijerph16122212

Armstrong, T., Wells, J., Boisvert, D. L., Lewis, R., Cooke, E. M., Woeckener, M., & Kavish, N. (2019). Skin conductance, heart rate and aggressive behavior type. *Biological Psychology*, *141*, 44–51. https://doi.org/10.1016/j.biopsycho.2018.12.012

Asbury, K., Dunn, J. F., Pike, A., & Plomin, R. (2003). Nonshared environmental influences on individual differences in early behavioral development: A monozygotic twin differences study. *Child Development*, *74*(3), 933–943. https://doi.org/10.1111/1467-8624.00577

Asbury, K., Dunn, J. F., & Plomin, R. (2006). Birthweight-discordance and differences in early parenting relate to monozygotic twin differences in behaviour problems and academic achievement at age 7. *Developmental Science*, *9*(2), F22–F31. https://doi.org/10.1111/j.1467-7687.2006.00469.x

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Averdijk, M., Ribeaud, D., & Eisner, M. (2018). The long-term effects of out-of-home placement in late adolescence: A propensity score matching analysis among Swiss youths. *Longitudinal and Life Course Studies*, *9*(1), 30–57. https://doi.org/10.14301/llcs.v9i1.450

Baadsgaard, M., & Quitzau, J. (2011). Danish registers on personal income and transfer payments. *Scandinavian Journal of Public Health*, *39*(7_suppl), 103–105. https://doi.org/10.1177/1403494811405098

Baker, L. A., Tuvblad, C., Reynolds, C., Zheng, M., Lozano, D. I., & Raine, A. (2009). Resting heart rate and the development of antisocial behavior from age 9 to 14: Genetic and environmental influences. *Development and Psychopathology*, *21*(3), 939–960. https://doi.org/10.1017/s0954579409000509

Baldwin, J. R., Reuben, A., Newbury, J. B., & Danese, A. (2019). Agreement Between Prospective and Retrospective Measures of Childhood Maltreatment: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *76*(6), 584–593. https://doi.org/10.1001/jamapsychiatry.2019.0097

Baldwin, J. R., Wang, B., Karwatowska, L., Schoeler, T., Tsaligopoulou, A., Munafò, M. R., & Pingault, J.-B. (2023). Childhood Maltreatment and Mental Health Problems: A Systematic Review and Meta-Analysis of Quasi-Experimental Studies. *American Journal of Psychiatry*, *180*(2), 117–126. https://doi.org/10.1176/appi.ajp.20220174

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development.* Holt, Rinehart & Winston.

Bann, D., Courtin, E., Davies, N. M., & Wright, L. (2024). Dialling back 'impact' claims: Researchers should not be compelled to make policy claims based on single studies. *International Journal of Epidemiology*, *53*(1), dyad181. https://doi.org/10.1093/ije/dyad181

Barker, D. J. P. (1990). The fetal and infant origins of adult disease. *BMJ*, *301*(6761), 1111–1111. https://doi.org/10.1136/bmj.301.6761.1111

Barker, E. D., & Maughan, B. (2009). Differentiating Early-Onset Persistent Versus Childhood-Limited Conduct Problem Youth. *American Journal of Psychiatry*, *166*(8), 900–908. https://doi.org/10.1176/appi.ajp.2009.08121770

Barnett, M. A., & Scaramella, L. V. (2013). Mothers' parenting and child sex differences in behavior problems among African American preschoolers. *Journal of Family Psychology*, *27*(5), 773–783. https://doi.org/10.1037/a0033792

Bärnighausen, T., Oldenburg, C., Tugwell, P., Bommer, C., Ebert, C., Barreto, M., Djimeu, E., Haber, N., Waddington, H., Rockers, P., Sianesi, B., Bor, J., Fink, G., Valentine, J., Tanner, J., Stanley, T., Sierra, E., Tchetgen, E. T., Atun, R., & Vollmer, S. (2017). Quasi-experimental study designs series—Paper 7: Assessing the assumptions. *Journal of Clinical Epidemiology*, *89*, 53–66. https://doi.org/10.1016/j.jclinepi.2017.02.017

Bärnighausen, T., Røttingen, J.-A., Rockers, P., Shemilt, I., & Tugwell, P. (2017). Quasi-experimental study designs series—Paper 1: Introduction: Two historical lineages. *Journal of Clinical Epidemiology*, *89*, 4–11. https://doi.org/10.1016/j.jclinepi.2017.02.020

Bärnighausen, T., Tugwell, P., Røttingen, J.-A., Shemilt, I., Rockers, P., Geldsetzer, P., Lavis, J., Grimshaw, J., Daniels, K., Brown, A., Bor, J., Tanner, J., Rashidian, A., Barreto, M., Vollmer, S., & Atun, R. (2017). Quasi-experimental study designs series—Paper 4: Uses and value. *Journal of Clinical Epidemiology*, *89*, 21–29. https://doi.org/10.1016/j.jclinepi.2017.03.012

Beach, S. R. H., Brody, G. H., Lei, M. K., Gibbons, F. X., Gerrard, M., Simons, R. L., Cutrona, C. E., & Philibert, R. A. (2013). Impact of child sex abuse on adult psychopathology: A genetically and epigenetically informed investigation. *Journal of Family Psychology*, *27*(1), 3–11. https://doi.org/10.1037/a0031459

Beauchaine, T. P., Bell, Z., Knapton, E., McDonough-Caplan, H., Shader, T., & Zisner, A. (2019). Respiratory sinus arrhythmia reactivity across empirically based structural dimensions of psychopathology: A meta-analysis. *Psychophysiology*, *56*(5), e13329. https://doi.org/10.1111/psyp.13329

Beauchaine, T. P., & Thayer, J. F. (2015). Heart rate variability as a transdiagnostic biomarker of psychopathology. *International Journal of Psychophysiology*, *98*(2), 338–350. https://doi.org/10.1016/j.ijpsycho.2015.08.004

Beaver, K. M., Connolly, E. J., Nedelec, J. L., & Schwartz, J. A. (2018). On the Genetic and Genomic Basis of Aggression, Violence, and Antisocial Behavior. In *Oxford Handbook of Evolution, Biology, and Society*.

Beck, A. T. (2008). The Evolution of the Cognitive Model of Depression and Its Neurobiological Correlates. *American Journal of Psychiatry*, *165*(8), 969–977. https://doi.org/10.1176/appi.ajp.2008.08050721

Becker, B. J., Aloe, A. M., Duvendack, M., Stanley, T. D., Valentine, J. C., Fretheim, A., & Tugwell, P. (2017). Quasi-experimental study designs series—Paper 10: Synthesizing evidence for effects collected from quasi-experimental studies presents surmountable challenges. *Journal of Clinical Epidemiology*, *89*, 84–91. https://doi.org/10.1016/j.jclinepi.2017.02.014

Bengtsson, J., Byberg, S., Carstensen, B., Stavola, B. L. D., Svensson, J., Jørgensen, M. E., & Rod, N. H. (2020). Accumulation of childhood adversities and type 1 diabetes risk: A register-based cohort study of all children born in Denmark between 1980 and 2015. *International Journal of Epidemiology*, *49*(5), 1604–1613. https://doi.org/10.1093/ije/dyaa138

Bengtsson, J., Dich, N., Rieckmann, A., & Rod, N. H. (2019). Cohort profile: The DANish LIFE course (DANLIFE) cohort, a prospective register-based cohort of all children born in Denmark since 1980. *BMJ Open*, *9*(9), e027217. https://doi.org/10.1136/bmjopen-2018-027217

Bengtsson, J., Rieckmann, A., Carstensen, B., Svensson, J., Jørgensen, M. E., & Rod, N. H. (2021). Trajectories of childhood adversity and type 1 diabetes: A nationwide study of 1 million children. *Diabetes Care*, *44*(3), 740–747. https://doi.org/10.2337/dc20-1130

Benichou, J., & Gail, M. H. (1990). Estimates of absolute cause-specific risk in cohort studies. *Biometrics*, *46*(3), 813–826. https://doi.org/10.2307/2532098

Ben-Shlomo, Y., & Kuh, D. (2002). A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology*, *31*, 285–293. https://doi.org/10.1093/ije/31.2.285

Berens, A. E., Jensen, S. K. G., & Nelson, C. A. (2017). Biological embedding of childhood adversity: From physiological mechanisms to clinical implications. *BMC Medicine*, *15*(1), 135. https://doi.org/10.1186/s12916-017-0895-4

Berg, L., Rostila, M., Arat, A., & Hjern, A. (2019). Parental death during childhood and violent crime in late adolescence to early adulthood: A Swedish national cohort study. *Palgrave Communications*, *5*(1), 74. https://doi.org/10.1057/s41599-019-0285-y

Berger, L. M., Hill, J., & Waldfogel, J. (2005). Maternity leave, early maternal employment and child health and development in the US. *The Economic Journal*, *115*(501), F29–F47. https://doi.org/10.1111/j.0013-0133.2005.00971.x

Bergstrøm, H., & Farrington, D. P. (2018). "The beat of my heart": The relationship between resting heart rate and psychopathy in a prospective longitudinal study. *Journal of Criminal Psychology*, *8*(4), 333–344. https://doi.org/10.1108/jcp-07-2018-0029

Besemer, S., Loeber, R., Hinshaw, S. P., & Pardini, D. A. (2016). Bidirectional associations between externalizing behavior problems and maladaptive parenting within parent-son dyads across childhood. *Journal of Abnormal Child Psychology*, *44*(7), 1387–1398. https://doi.org/10.1007/s10802-015-0124-6

Bevilacqua, L., Hale, D., Barker, E. D., & Viner, R. (2018). Conduct problems trajectories and psychosocial outcomes: A systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, *27*(10), 1239–1260. https://doi.org/10.1007/s00787-017-1053-4

Boisvert, D., & Wright, J. P. (2008). Nonshared environmental influences on sibling differences in externalizing problem behavior. *Criminal Justice and Behavior*, *35*(7), 863–878. https://doi.org/10.1177/0093854808318495

Book, A. S., Starzyk, K. B., & Quinsey, V. L. (2001). The relationship between testosterone and aggression: A meta-analysis. *Aggression and Violent Behavior*, *6*(6), 579–599. https://doi.org/10.1016/S1359-1789(00)00032-X

Bor, J., Moscoe, E., Mutevedzi, P., Newell, M.-L., & Bärnighausen, T. (2014). Regression discontinuity designs in epidemiology: Causal inference without randomized trials. *Epidemiology*, *25*(5), 729–737. https://doi.org/10.1097/ede.0000000000000138

Bornstein, M. H. (1989). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin*, *105*(2), 179–197. https://doi.org/10.1037/0033-2909.105.2.179

Boutwell, B. B., & Beaver, K. M. (2010). Maternal cigarette smoking during pregnancy and offspring externalizing behavioral problems: A propensity score matching analysis. *International Journal of Environmental Research and Public Health*, *7*(1), 146–163. https://doi.org/10.3390/ijerph7010146

Boutwell, B. B., Beaver, K. M., Gibson, C. L., & Ward, J. T. (2011). Prenatal exposure to cigarette smoke and childhood externalizing behavioral problems: A propensity score matching approach. *International Journal of Environmental Health Research*, *21*(4), 248–259. https://doi.org/10.1080/09603123.2010.544032

Bowden, J., Davey Smith, G., Haycock, P. C., & Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic Epidemiology*, *40*(4), 304–314. https://doi.org/10.1002/gepi.21965

Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N. A., & Thompson, J. R. (2016). Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: The role of the I2 statistic. *International Journal of Epidemiology*, *45*(6), 1961–1974. https://doi.org/10.1093/ije/dyw220

Bowden, J., Smith, G. D., & Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression.

International Journal of Epidemiology, *44*(2), 512–525. https://doi.org/10.1093/ije/dyv080

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*(2), 129–148. https://doi.org/10.1037/0003-066X.36.2.129

Boyle, M. H., Jenkins, J. M., Georgiades, K., Cairney, J., Duku, E., & Racine, Y. (2004). Differential-maternal parenting behavior: Estimating within- and between-family effects on children. *Child Development*, *75*(5), 1457–1476. https://doi.org/10.1111/j.1467-8624.2004.00751.x

Bradshaw, D., Hannigan, A., Creaven, A., & Muldoon, O. T. (2020). Longitudinal associations between parental incarceration and children's emotional and behavioural development: Results from a population cohort study. *Child: Care, Health and Development*, *46*(2), 195–202. https://doi.org/10.1111/cch.12732

Brandlistuen, R. E., Ystrom, E., Eberhard-Gran, M., Nulman, I., Koren, G., & Nordeng, H. (2015). Behavioural effects of fetal antidepressant exposure in a Norwegian cohort of discordant siblings. *International Journal of Epidemiology*, *44*(4), 1397–1407. https://doi.org/10.1093/ije/dyv030

Brandlistuen, R. E., Ystrom, E., Hernandez-Diaz, S., Skurtveit, S., Selmer, R., Handal, M., & Nordeng, H. (2017). Association of prenatal exposure to benzodiazepines and child internalizing problems: A sibling-controlled cohort study. *PLoS ONE*, *12*(7), e0181042. https://doi.org/10.1371/journal.pone.0181042

Brandlistuen, R. E., Ystrom, E., Nulman, I., Koren, G., & Nordeng, H. (2013). Prenatal paracetamol exposure and child neurodevelopment: A sibling-controlled cohort study. *International Journal of Epidemiology*, *42*(6), 1702–1713. https://doi.org/10.1093/ije/dyt183

Brewin, C. R., Andrews, B., & Gotlib, I. H. (1993). Psychopathology and Early Experience: A Reappraisal of Retrospective Reports. *Psychological Bulletin*, *113(1)*, 82–98. https://doi.org/10.1037/0033-2909.113.1.82

Briggs, E. C., Amaya-Jackson, L., Putnam, K. T., & Putnam, F. W. (2021). All adverse childhood experiences are not equal: The contribution of synergy to adverse childhood experience scores. *American Psychologist*, *76*(2), 243–252. https://doi.org/10.1037/amp0000768

Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie, G. Å., 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., … Davies, N. M. (2020). Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nature Communications*, *11*(1), 3519. https://doi.org/10.1038/s41467-020-17117-4

Brunborg, G. S., Mentzoni, R. A., & Frøyland, L. R. (2014). Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking, or conduct problems? *Journal of Behavioral Addictions*, *3*(1), 27–32. https://doi.org/10.1556/JBA.3.2014.002

Bubonya, M., Cobb-Clark, D., Christensen, D., Johnson, S., & Zubrick, S. (2019). The Great Recession and children's mental health in Australia. *International Journal of Environmental Research and Public Health*, *16*(4), 537. https://doi.org/10.3390/ijerph16040537

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., Duncan, L., Perry, J. R. B., Patterson, N., Robinson, E. B., Daly, M. J., Price, A. L., & Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, *47*(11), Article 11. https://doi.org/10.1038/ng.3406

Burgess, S., Butterworth, A., & Thompson, S. G. (2013). Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. *Genetic Epidemiology*, *37*(7), 658–665. https://doi.org/10.1002/gepi.21758

Burgess, S., Foley, C. N., Allara, E., Staley, J. R., & Howson, J. M. M. (2020). A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-019-14156-4

Burgess, S., Small, D. S., & Thompson, S. G. (2017). A review of instrumental variable estimators for Mendelian randomization. *Statistical Methods in Medical Research*, *26*(5), 2333–2355. https://doi.org/10.1177/0962280215597579

Burgess, S., Timpson, N. J., Ebrahim, S., & Smith, G. D. (2015). Mendelian randomization: Where are we now and where are we going? *International Journal of Epidemiology*, *44*(2), 379–388. https://doi.org/10.1093/ije/dyv108

Burt, S. A., Barnes, A. R., McGue, M., & Iacono, W. G. (2008). Parental divorce and adolescent delinquency: Ruling out the impact of common genes. *Developmental Psychology*, *44*(6), 1668–1677. https://doi.org/10.1037/a0013477

Burt, S. A., Donnellan, M. B., Humbad, M. N., Hicks, B. M., McGue, M., & Iacono, W. G. (2010). Does marriage inhibit antisocial behavior?: An examination of selection vs causation via a longitudinal twin design. *Archives of General Psychiatry*, *67*(12), 1309–1315. https://doi.org/10.1001/archgenpsychiatry.2010.159

Burt, S. A., Hyde, L. W., Frick, P. J., Jaffee, S. R., Shaw, D. S., & Tremblay, R. (2018). Commentary: Childhood conduct problems are a public health crisis and

require resources: a commentary on Rivenbark et al. (2018). *Journal of Child Psychology and Psychiatry*, *59*(6), 711–713. https://doi.org/10.1111/jcpp.12930

Burt, S. A., McGue, M., & Iacono, W. G. (2009). Nonshared environmental mediation of the association between deviant peer affiliation and adolescent externalizing behaviors over time: Results from a cross-lagged monozygotic twin differences design. *Developmental Psychology*, *45*(6), 1752–1760. https://doi.org/10.1037/a0016687

Burt, S. A., McGue, M., Krueger, R. F., & Iacono, W. G. (2005). How are parent–child conflict and childhood externalizing symptoms related over time? Results from a genetically informative cross-lagged study. *Development and Psychopathology*, *17*(1), 145–165. https://doi.org/10.1017/S095457940505008X

Cadoret, R. J. (1995). Genetic-environmental interaction in the genesis of aggressivity and conduct disorders. *Archives of General Psychiatry*, *52*(11), 916–924. https://doi.org/10.1001/archpsyc.1995.03950230030006

Cadoret, R. J., & Stewart, M. A. (1991). An adoption study of attention deficit/hyperactivity/aggression and their relationship to adult antisocial personality. *Comprehensive Psychiatry*, *32*(1), 73–82. https://doi.org/10.1016/0010-440X(91)90072-K

Campbell, D. T. (Donald T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Houghton Mifflin Company.

Carter Leno, V., Chandler, S., White, P., Yorke, I., Charman, T., Jones, C. R. G., Happé, F., Baird, G., Pickles, A., & Simonoff, E. (2021). Associations between theory of mind and conduct problems in autistic and nonautistic youth. *Autism Research*, *14*(2), 276–288. https://doi.org/10.1002/aur.2346

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. https://doi.org/10.1037/0012-1649.40.2.149

Ceccarelli, C., Prina, E., Muneghina, O., Jordans, M., Barker, E., Miller, K., Singh, R., Acarturk, C., Sorsdhal, K., Cuijpers, P., Lund, C., Barbui, C., & Purgato, M. (2022). Adverse childhood experiences and global mental health: Avenues to reduce the burden of child and adolescent mental disorders. *Epidemiology and Psychiatric Sciences*, *31*(e75), 1–7. https://doi.org/10.1017/S2045796022000580

Cecil, C. A. M., Barker, E. D., Jaffee, S. R., & Viding, E. (2012). Association between maladaptive parenting and child self-control over time: Cross-lagged study

using a monozygotic twin difference design. *British Journal of Psychiatry*, *201*(4), 291–297. https://doi.org/10.1192/bjp.bp.111.107581

Chao, M., Li, X., & McGue, M. (2017). The causal role of alcohol use in adolescent externalizing and internalizing problems: A Mendelian randomization study. *Alcoholism: Clinical and Experimental Research*, *41*(11), 1953–1960. https://doi.org/10.1111/acer.13493

Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases*, *40*(5), 373–383. https://doi.org/10.1016/0021-9681(87)90171-8

Cole, S. R., Platt, R. W., Schisterman, E. F., Chu, H., Westreich, D., Richardson, D., & Poole, C. (2010). Illustrating bias due to conditioning on a collider. *International Journal of Epidemiology*, *39*(2), 417–420. https://doi.org/10.1093/ije/dyp334

Colman, I., Murray, J., Abbott, R. A., Maughan, B., Kuh, D., Croudace, T. J., & Jones, P. B. (2009). Outcomes of conduct problems in adolescence: 40 year follow-up of national cohort. *BMJ*, *338*, a2981. https://doi.org/10.1136/bmj.a2981

Cooke, J. E., Deneault, A.-A., Devereux, C., Eirich, R., Fearon, R. M. P., & Madigan, S. (2022). Parental sensitivity and child behavioral problems: A meta-analytic review. *Child Development*, *93*(5), 1231–1248. https://doi.org/10.1111/cdev.13764

Copeland, K. T., Checkoway, H., McMichael, A. J., & Holbrook, R. H. (1977). Bias due to misclassification in the estimation of relative risk. *American Journal of Epidemiology*, *105*(5), 488–495. https://doi.org/10.1093/oxfordjournals.aje.a112408

Copeland, W. E., Shanahan, L., Costello, E. J., & Angold, A. (2009). Childhood and adolescent psychiatric disorders as predictors of young adult disorders. *Archives of General Psychiatry*, *66*(7), 764–772. https://doi.org/10.1001/archgenpsychiatry.2009.85

Copp, J. E., Giordano, P. C., Manning, W. D., & Longmore, M. A. (2018). Parental incarceration and child well-being: Conceptual and practical concerns regarding the use of propensity scores. *Socius: Sociological Research for a Dynamic World*, *4*, 1–12. https://doi.org/10.1177/2378023118779306

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *JAMA*, *290*(15), 2023–2029. https://doi.org/10.1001/jama.290.15.2023

Costello, E. J., Erkanli, A., Copeland, W., & Angold, A. (2010). Association of family income supplements in adolescence with development of psychiatric and

substance use disorders in adulthood among an American Indian population. *JAMA*, *303*(19), 1954–1960. https://doi.org/10.1001/jama.2010.621

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187–220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

Cree, R. A., Liu, C., Gueorguieva, R., Neiderhiser, J. M., Leve, L. D., Connell, C. M., Shaw, D. S., Natsuaki, M. N., Ganiban, J. M., Beekman, C., Smith, M. V., & Reiss, D. (2021). Using an adoption design to test genetically based differences in risk for child behavior problems in response to home environmental influences. *Development and Psychopathology*, *33*(4), 1229–1247. https://doi.org/10.1017/s0954579420000450

Crosby, D. A., Dowsett, C. J., Gennetian, L. A., & Huston, A. C. (2010). A tale of two methods: Comparing regression and instrumental variables estimates of the effects of preschool child care type on the subsequent externalizing behavior of children in low-income families. *Developmental Psychology*, *46*(5), 1030–1048. https://doi.org/10.1037/a0020384

Danese, A., Caspi, A., Williams, B., Ambler, A., Sugden, K., Mika, J., Werts, H., Freeman, J., Pariante, C., Moffitt, T., & Arseneault, L. (2011). Biological embedding of stress through inflammation processes in childhood. *Molecular Psychiatry*, *16*(3), 244–246. https://doi.org/10.1038/mp.2010.5

Danese, A., Pariante, C. M., Caspi, A., Taylor, A., & Poulton, R. (2007). Childhood maltreatment predicts adult inflammation in a life-course study. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(4), 1319–1324. https://doi.org/10.1073/pnas.0610362104

Daniel, R. M., Cousens, S. N., Stavola, B. L. D., Kenward, M. G., & Sterne, J. A. C. (2013). Methods for dealing with time-dependent confounding. *Statistics in Medicine*, *32*(9), 1584–1618. https://doi.org/10.1002/sim.5686

Davey Smith, G., & Ebrahim, S. (2003). 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Davey Smith, G., & Hemani, G. (2014). Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics*, *23*(R1), R89–R98. https://doi.org/10.1093/hmg/ddu328

Davies, N. M., Holmes, M. V., & Smith, G. D. (2017). Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ*, *362*, k601. https://doi.org/10.1136/bmj.k601

De Brito, S. A., Forth, A. E., Baskin-Sommers, A. R., Brazil, I. A., Kimonis, E. R., Pardini, D., Frick, P. J., Blair, R. J. R., & Viding, E. (2021). Psychopathy. *Nature Reviews Disease Primers*, *7*(1), 49. https://doi.org/10.1038/s41572-021-00282-1

de Geus, E., Kupper, N., Boomsma, D. I., & Snieder, H. (2007). Bivariate Genetic Modeling of Cardiovascular Stress Reactivity: Does Stress Uncover Genetic Variance? *Psychosomatic Medicine*, *69*(4), 356–364. https://doi.org/10.1097/psy.0b013e318049cc2d

de Looff, P. C., Cornet, L. J. M., Kogel, C. H. de, Fernández-Castilla, B., Embregts, P. J. C. M., Didden, R., & Nijman, H. L. I. (2022). Heart rate and skin conductance associations with physical aggression, psychopathy, antisocial personality disorder and conduct disorder: An updated meta-analysis. *Neuroscience & Biobehavioral Reviews*, *132*, 553–582. https://doi.org/10.1016/j.neubiorev.2021.11.003

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*, *37*(5), 637–652. https://doi.org/10.1007/s10802-009-9307-3

De Stavola, B. L., Herle, M., & Pickles, A. (2022). Framing causal questions in life course epidemiology. *Annual Review of Statistics and Its Application*, *9*(1), 223–248. https://doi.org/10.1146/annurev-statistics-040120-024748

de Vries, T. R., Arends, I., Hulvej Rod, N., Oldehinkel, A. J., & Bültmann, U. (2022). Proposing network analysis for early life adversity: An application on life event data. *Social Science & Medicine*, *296*, 114784. https://doi.org/10.1016/j.socscimed.2022.114784

Deater-Deckard, K., & Petrill, S. A. (2004). Parent–child dyadic mutuality and child behavior problems: An investigation of gene–environment processes. *Journal of Child Psychology and Psychiatry*, *45*(6), 1171–1179. https://doi.org/10.1111/j.1469-7610.2004.00309.x

Dee, T. S., & Sievertsen, H. H. (2018). The gift of time? School starting age and mental health. *Health Economics*, *27*(5), 781–802. https://doi.org/10.1002/hec.3638

Derzon, J. H. (2010). The correspondence of family features with problem, aggressive, criminal, and violent behavior: A meta-analysis. *Journal of Experimental Criminology*, *6*(3), 263–292. https://doi.org/10.1007/s11292-010-9098-0

Dinwiddie, S., Heath, A. C., Dunne, M. P., Bucholz, K. K., Madden, P. A. F., Slutske, W. S., Bierut, L. J., Statham, D. B., & Martin, N. G. (2000). Early sexual abuse and lifetime psychopathology: A co-twin–control study. *Psychological Medicine*, *30*(1), 41–52. https://doi.org/10.1017/S0033291799001373

Dodge, K. A., Pettit, G. S., Bates, J. E., & Valente, E. (1995). Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. *Journal of Abnormal Psychology*, *104*(4), 632–643. https://doi.org/10.1037/0021-843X.104.4.632

Doi, S., Fujiwara, T., Isumi, A., Ochi, M., & Kato, T. (2018). Relationship between leaving children at home alone and their mental health: Results from the A-CHILD study in Japan. *Frontiers in Psychiatry*, *9*, 192. https://doi.org/10.3389/fpsyt.2018.00192

Doidge, J. C., Edwards, B., Higgins, D. J., & Segal, L. (2017). Adverse childhood experiences, non-response and loss to follow-up: Findings from a prospective birth cohort and recommendations for addressing missing data. *Longitudinal and Life Course Studies*, *8*(4), Article 4. https://doi.org/10.14301/llcs.v8i4.414

Dong, M., Anda, R. F., Felitti, V. J., Dube, S. R., Williamson, D. F., Thompson, T. J., Loo, C. M., & Giles, W. H. (2004). The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Child Abuse & Neglect*, *28*(7), 771–784. https://doi.org/10.1016/j.chiabu.2004.01.008

D'Onofrio, B. M., Goodnight, J. A., Hulle, C. A. V., Rodgers, J. L., Rathouz, P. J., Waldman, I. D., & Lahey, B. B. (2009a). A quasi-experimental analysis of the association between family income and offspring conduct problems. *Journal of Abnormal Child Psychology*, *37*(3), 415–429. https://doi.org/10.1007/s10802-008-9280-2

D'Onofrio, B. M., Goodnight, J. A., Hulle, C. A. V., Rodgers, J. L., Rathouz, P. J., Waldman, I. D., & Lahey, B. B. (2009b). Maternal age at childbirth and offspring disruptive behaviors: Testing the causal hypothesis. *Journal of Child Psychology and Psychiatry*, *50*(8), 1018–1028. https://doi.org/10.1111/j.1469-7610.2009.02068.x

D'Onofrio, B. M., Hulle, C. A. V., Waldman, I. D., Rodgers, J. L., Harden, K. P., Rathouz, P. J., & Lahey, B. B. (2008). Smoking during pregnancy and offspring externalizing problems: An exploration of genetic and environmental confounds. *Development and Psychopathology*, *20*(1), 139–164. https://doi.org/10.1017/S0954579408000072

D'Onofrio, B. M., Slutske, W. S., Turkheimer, E., Emery, R. E., Harden, K. P., Heath, A. C., Madden, P. A. F., & Martin, N. G. (2007). Intergenerational transmission of childhood conduct problems: A Children of Twins Study. *Archives of General Psychiatry*, *64*(7), 820–829. https://doi.org/10.1001/archpsyc.64.7.820

D'Onofrio, B. M., Turkheimer, E., Emery, R. E., Slutske, W. S., Heath, A. C., Madden, P. A., & Martin, N. G. (2005). A genetically informed study of marital instability

and its association with offspring psychopathology. *Journal of Abnormal Psychology*, *114*(4), 570–586. https://doi.org/10.1037/0021-843X.114.4.570

D'Onofrio, B. M., Van Hulle, C. A., Goodnight, J. A., Rathouz, P. J., & Lahey, B. B. (2012). Is maternal smoking during pregnancy a causal environmental risk factor for adolescent antisocial behavior? Testing etiological theories and assumptions. *Psychological Medicine*, *42*(7), 1535–1545. https://doi.org/10.1017/S0033291711002443

D'Onofrio, B. M., Van Hulle, C. A., Waldman, I. D., Rodgers, J. L., Rathouz, P. J., & Lahey, B. B. (2007). Causal inferences regarding prenatal alcohol exposure and childhood externalizing problems. *Archives of General Psychiatry*, *64*(11), 1296–1304. https://doi.org/10.1001/archpsyc.64.11.1296

Dunifon, R., Kalil, A., & Danziger, S. K. (2003). Maternal work behavior under welfare reform: How does the transition from welfare to work affect child development? *Children and Youth Services Review*, *25*(1–2), 55–82. https://doi.org/10.1016/S0190-7409(02)00266-9

Duyme, M. (1990). Antisocial behaviour and postnatal environment: A French adoption study. *Journal of Child Psychology and Psychiatry*, *31*(5), 699–710. https://doi.org/10.1111/j.1469-7610.1990.tb00811.x

Edwards, B., & Yu, M. (2018). The influence of child care on the behavior problems of children of teenage mothers. *Children and Youth Services Review*, *94*, 96–104. https://doi.org/10.1016/j.childyouth.2018.09.029

Egeland, B., Yates, T., Appleyard, K., & Van Dulmen, M. (2002). The Long-Term Consequences of Maltreatment in the Early Years: A Developmental Pathway Model to Antisocial Behavior. *Children's Services*, *5*(4), 249–260. https://doi.org/10.1207/S15326918CS0504_2

Ekblad, M., Lehtonen, L., Korkeila, J., & Gissler, M. (2017). Maternal smoking during pregnancy and the risk of psychiatric morbidity in singleton sibling pairs. *Nicotine & Tobacco Research*, *19*(5), 597–604. https://doi.org/10.1093/ntr/ntx001

Ekblad, M., Rolan, E., Marceau, K., Palmer, R. H. C., Todorov, A., Todorov, A. A., Heath, A. C., & Knopik, V. S. (2020). Disruptive behavior in siblings discordant for exposure to maternal smoking during pregnancy: A multi-rater approach. *Nicotine & Tobacco Research*, *22*(8), 1330–1338. https://doi.org/10.1093/ntr/ntz214

El-Khoury, F., Rieckmann, A., Bengtsson, J., Melchior, M., & Rod, N. H. (2021). Childhood adversity trajectories and PTSD in young adulthood: A nationwide Danish register-based cohort study of more than one million individuals.

Journal of Psychiatric Research, *136*, 274–280. https://doi.org/10.1016/j.jpsychires.2021.02.034

Ellingson, J. M., Goodnight, J. A., Van Hulle, C. A., Waldman, I. D., & D'Onofrio, B. M. (2014). A sibling-comparison study of smoking during pregnancy and childhood psychological traits. *Behavior Genetics*, *44*(1), 25–35. https://doi.org/10.1007/s10519-013-9618-6

Ellis, L. C., Berg-Nielsen, T. S., Lydersen, S., & Wichstrøm, L. (2012). Smoking during pregnancy and psychiatric disorders in preschoolers. *European Child & Adolescent Psychiatry*, *21*(11), 635–644. https://doi.org/10.1007/s00787-012-0300-y

Ellis, L., & Walsh, A. (2003). Crime, delinquency and intelligence: A review of the worldwide literature. In H. Nyborg (Ed.), *The Scientific Study of General Intelligence* (pp. 343–365). Elsevier. https://doi.org/10.1016/B978-008043793-4/50054-4

Emery, C. R. (2011). Controlling for selection effects in the relationship between child behavior problems and exposure to intimate partner violence. *Journal of Interpersonal Violence*, *26*(8), 1541–1558. https://doi.org/10.1177/0886260510370597

Eppinga, R. N., Hagemeijer, Y., Burgess, S., Hinds, D. A., Stefansson, K., Gudbjartsson, D. F., Veldhuisen, D. J. van, Munroe, P. B., Verweij, N., & Harst, P. van der. (2016). Identification of genomic loci associated with resting heart rate and shared genetic predictors with all-cause mortality. *Nature Genetics*, *48*(12), 1557–1563. https://doi.org/10.1038/ng.3708

Erskine, H. E., Ferrari, A. J., Polanczyk, G. V., Moffitt, T. E., Murray, C. J. L., Vos, T., Whiteford, H. A., & Scott, J. G. (2014). The global burden of conduct disorder and attention-deficit/hyperactivity disorder in 2010. *Journal of Child Psychology and Psychiatry*, *55*(4), 328–336. https://doi.org/10.1111/jcpp.12186

Estabrook, R., Massey, S. H., Clark, C. A. C., Burns, J. L., Mustanski, B. S., Cook, E. H., O'Brien, T. C., Makowski, B., Espy, K. A., & Wakschlag, L. S. (2016). Separating family-level and direct exposure effects of smoking during pregnancy on offspring externalizing symptoms: Bridging the behavior genetic and behavior teratologic divide. *Behavior Genetics*, *46*(3), 389–402. https://doi.org/10.1007/s10519-015-9762-2

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin*, *139*(6), 1342–1396. https://doi.org/10.1037/a0031808

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-Based Psychosocial Treatments for Children and Adolescents With Disruptive Behavior. *Journal of*

Clinical Child & Adolescent Psychology, *37*(1), 215–237. https://doi.org/10.1080/15374410701820117

Fagundes, C. P., Glaser, R., & Kiecolt-Glaser, J. K. (2013). Stressful early life experiences and immune dysregulation across the lifespan. *Brain, Behavior, and Immunity*, *27*, 8–12. https://doi.org/10.1016/j.bbi.2012.06.014

Fairchild, G., Hawes, D. J., Frick, P. J., Copeland, W. E., Odgers, C. L., Franke, B., Freitag, C. M., & Brito, S. A. D. (2019). Conduct disorder. *Nature Reviews Disease Primers*, *5*(1), 43. https://doi.org/10.1038/s41572-019-0095-y

Fanti, K. A. (2018). Understanding heterogeneity in conduct disorder: A review of psychophysiological studies. *Neuroscience & Biobehavioral Reviews*, *91*, 4–20. https://doi.org/10.1016/j.neubiorev.2016.09.022

Fanti, K. A., Eisenbarth, H., Goble, P., Demetriou, C., Kyranides, M. N., Goodwin, D., Zhang, J., Bobak, B., & Cortese, S. (2019). Psychophysiological activity and reactivity in children and adolescents with conduct problems: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *100*, 98–107. https://doi.org/10.1016/j.neubiorev.2019.02.016

Faraone, S. V., & Larsson, H. (2019). Genetics of attention deficit hyperactivity disorder. *Molecular Psychiatry*, *24*(4), 562–575. https://doi.org/10.1038/s41380-018-0070-0

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine*, *14*(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1994). Parental Separation, Adolescent Psychopathology, and Problem Behaviors. *Journal of the American Academy of Child & Adolescent Psychiatry*, *33*(8), 1122–1133. https://doi.org/10.1097/00004583-199410000-00008

Fine, J. P., & Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, *94*(446), 496–509. https://doi.org/10.1080/01621459.1999.10474144

Fitzsimons, E., & Villadsen, A. (2019). Father departure and children's mental health: How does timing matter? *Social Science & Medicine*, *222*, 349–358. https://doi.org/10.1016/j.socscimed.2018.11.008

Francis, E. R., Tsaligopoulou, A., Stock, S. E., Pingault, J., & Baldwin, J. R. (2023). Subjective and objective experiences of childhood adversity: A meta-analysis of their agreement and relationships with psychopathology. *Journal of Child Psychology and Psychiatry*, jcpp.13803. https://doi.org/10.1111/jcpp.13803

Frick, P. J., Cornell, A. H., Bodin, S. D., Dane, H. E., Barry, C. T., & Loney, B. R. (2003). Callous-unemotional traits and developmental pathways to severe conduct problems. *Developmental Psychology*, *39*(2), 246–260. https://doi.org/10.1037/0012-1649.39.2.246

Frick, P. J., & Ellis, M. (1999). Callous-unemotional traits and subtypes of conduct disorder. *Clinical Child and Family Psychology Review*, *2*(3), 149–168. https://doi.org/10.1023/a:1021803005547

Frick, P. J., Lahey, B. B., Loeber, R., Stouthamer-Loeber, M., Christ, M. A. G., & Hanson, K. (1992). Familial risk factors to oppositional defiant disorder and conduct disorder: Parental psychopathology and maternal parenting. *Journal of Consulting and Clinical Psychology*, *60*(1), 49–55. https://doi.org/10.1037/0022-006X.60.1.49

Frisell, T. (2020). Sibling comparison designs, are they worth the effort? *American Journal of Epidemiology*, *190*(5), kwaa183. https://doi.org/10.1093/aje/kwaa183

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Funk, M. J., Westreich, D., Wiesen, C., Stürmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly Robust Estimation of Causal Effects. *American Journal of Epidemiology*, *173*(7), 761–767. https://doi.org/10.1093/aje/kwq439

Galán, C. A., Choe, D. E., Forbes, E. E., & Shaw, D. S. (2017). Interactions between empathy and resting heart rate in early adolescence predict violent behavior in late adolescence and early adulthood. *Journal of Child Psychology and Psychiatry*, *58*(12), 1370–1380. https://doi.org/10.1111/jcpp.12771

Ganiban, J. M., Liu, C., Zappaterra, L., An, S., Natsuaki, M. N., Neiderhiser, J. M., Reiss, D., Shaw, D. S., & Leve, L. D. (2021). Gene × Environment Interactions in the Development of Preschool Effortful Control, and Its Implications for Childhood Externalizing Behavior. *Behavior Genetics*, *51*(5), 448–462. https://doi.org/10.1007/s10519-021-10073-9

Gao, Y., Raine, A., Venables, P. H., Dawson, M. E., & Mednick, S. A. (2010). Association of Poor Childhood Fear Conditioning and Adult Crime. *American Journal of Psychiatry*, *167*(1), 56–60. https://doi.org/10.1176/appi.ajp.2009.09040499

Gardner, F., Leijten, P., Melendez-Torres, G. J., Landau, S., Harris, V., Mann, J., Beecham, J., Hutchings, J., & Scott, S. (2019). The Earlier the Better? Individual Participant Data and Traditional Meta-analysis of Age Effects of Parenting Interventions. *Child Development*, *90*(1), 7–19. https://doi.org/10.1111/cdev.13138

Gaysina, D., Fergusson, D. M., Leve, L. D., Horwood, J., Reiss, D., Shaw, D. S., Elam, K. K., Natsuaki, M. N., Neiderhiser, J. M., & Harold, G. T. (2013). Maternal smoking during pregnancy and offspring conduct problems: Evidence from three independent genetically-sensitive research designs. *JAMA Psychiatry*, *70*(9), 956–963. https://doi.org/10.1001/jamapsychiatry.2013.127

Gerds, T., Ohlendorff, J., & Ozenne, B. (2023). *riskRegression: Risk regression models and prediction scores for survival analysis with competing risks*. (R package version 2023.03.22) [Computer software].

Gershoff, E. T., Sattler, K. M. P., & Ansari, A. (2018). Strengthening causal estimates for links between spanking and children's externalizing behavior problems. *Psychological Science*, *29*(1), 110–120. https://doi.org/10.1177/0956797617729816

Gilbert, R., Widom, C. S., Browne, K., Fergusson, D., Webb, E., & Janson, S. (2009). Burden and consequences of child maltreatment in high-income countries. *The Lancet*, *373*(9657), 68–81. https://doi.org/10.1016/S0140-6736(08)61706-7

Gilman, S. E., Gardener, H., & Buka, S. L. (2008). Maternal smoking during pregnancy and children's cognitive and physical development: A causal risk factor? *American Journal of Epidemiology*, *168*(5), 522–531. https://doi.org/10.1093/aje/kwn175

Girard, L.-C., Doyle, O., & Tremblay, R. E. (2018). Breastfeeding and externalising problems: A quasi-experimental design with a national cohort. *European Child & Adolescent Psychiatry*, *27*(7), 877–884. https://doi.org/10.1007/s00787-017-1085-9

Girard, L.-C., & Farkas, C. (2019). Breastfeeding and behavioural problems: Propensity score matching with a national cohort of infants in Chile. *BMJ Open*, *9*(2), e025058. https://doi.org/10.1136/bmjopen-2018-025058

Gjerde, L. C., Eilertsen, E. M., Eley, T. C., McAdams, T. A., Reichborn-Kjennerud, T., Røysamb, E., & Ystrom, E. (2020). Maternal perinatal and concurrent anxiety and mental health problems in early childhood: A sibling-comparison study. *Child Development*, *91*(2), 456–470. https://doi.org/10.1111/cdev.13192

Gjerde, L. C., Eilertsen, E. M., Reichborn-Kjennerud, T., McAdams, T. A., Zachrisson, H. D., Zambrana, I. M., Røysamb, E., Kendler, K. S., & Ystrom, E. (2017). Maternal perinatal and concurrent depressive symptoms and child behavior problems: A sibling comparison study. *Journal of Child Psychology and Psychiatry*, *58*(7), 779–786. https://doi.org/10.1111/jcpp.12704

Glanville, J., Eyers, J., Jones, A. M., Shemilt, I., Wang, G., Johansen, M., Fiander, M., & Rothstein, H. (2017). Quasi-experimental study designs series—paper 8: Identifying quasi-experimental studies to inform systematic reviews. *Journal*

of Clinical Epidemiology, *89*, 67–76. https://doi.org/10.1016/j.jclinepi.2017.02.018

Glover, M. B., Mullineaux, P. Y., Deater-Deckard, K., & Petrill, S. A. (2010). Parents' feelings towards their adoptive and non-adoptive children. *Infant and Child Development*, *19*(3), 238–251. https://doi.org/10.1002/icd.664

Goetghebeur, E., Cessie, S. le, Stavola, B. D., Moodie, E. E., Waernbaum, I., & initiative, "on behalf of" the topic group Causal Inference (TG7) of the STRATOS. (2020). Formulating causal questions and principled statistical answers. *Statistics in Medicine*, *39*(30), 4922–4948. https://doi.org/10.1002/sim.8741

Goldberg, J. S., & Carlson, M. J. (2014). Parents' relationship quality and children's behavior in stable married and cohabiting families: Parents' relationship quality and children's behavior. *Journal of Marriage and Family*, *76*(4), 762–777. https://doi.org/10.1111/jomf.12120

Gomajee, R., El-Khoury, F., Côté, S., Van Der Waerden, J., Pryor, L., & Melchior, M. (2018). Early childcare type predicts children's emotional and behavioural trajectories into middle childhood. Data from the EDEN mother–child cohort study. *Journal of Epidemiology and Community Health*, *72*(11), 1033–1043. https://doi.org/10.1136/jech-2017-210393

Goodnight, J. A., Donahue, K. L., Waldman, I. D., Van Hulle, C. A., Rathouz, P. J., Lahey, B. B., & D'Onofrio, B. M. (2016). Genetic and environmental contributions to associations between infant fussy temperament and antisocial behavior in childhood and adolescence. *Behavior Genetics*, *46*(5), 680–692. https://doi.org/10.1007/s10519-016-9794-2

Gottschalk, L. A., & Gleser, G. C. (1979). *The measurement of psychological states through the content analysis of verbal behavior*. Univ of California Press.

Grabow, A. P., Khurana, A., Natsuaki, M. N., Neiderhiser, J. M., Harold, G. T., Shaw, D. S., Ganiban, J. M., Reiss, D., & Leve, L. D. (2017). Using an adoption–biological family design to examine associations between maternal trauma, maternal depressive symptoms, and child internalizing and externalizing behaviors. *Development and Psychopathology*, *29*(5), 1707–1720. https://doi.org/10.1017/S0954579417001341

Green, H., Britain, G., Britain, G., & Scotland. (2005). *Mental health of children and young people in Great Britain, 2004*. Palgrave Macmillan. notion://www.notion.so/Green-et-al-2005-b75e630fc0fb4234aed67f215457c0cb

Greenland, S., Pearl, J., & Robins, J. M. (1999). Causal diagrams for epidemiologic research. *Epidemiology*, 37–48. https://doi.org/10.1097/00001648-199901000-00008

Gunasekara, F. I., Richardson, K., Carter, K., & Blakely, T. (2014). Fixed effects analysis of repeated measures data. *International Journal of Epidemiology*, *43*(1), 264–269. https://doi.org/10.1093/ije/dyt221

Haan, E., Kirsten E Westmoreland, Kirsten E Westmoreland, Schellhas, L., Hannah M Sallis, Hannah M Sallis, Taylor, G. M. J., Luisa Zuccolo, Luisa Zuccolo, & Munafò, M. R. (2022). Prenatal smoking, alcohol and caffeine exposure and offspring externalizing disorders: A systematic review and meta-analysis. *Addiction*, 2602–2613. https://doi.org/10.1111/add.15858

Haber, J. R., Bucholz, K. K., Jacob, T., Grant, J. D., Scherrer, J. F., Sartor, C. E., Duncan, A. E., & Heath, A. (2010). Effect of paternal alcohol and drug dependence on offspring conduct disorder: Gene-environment interplay. *Journal of Studies on Alcohol and Drugs*, *71*(5), 652–663. https://doi.org/10.15288/jsad.2010.71.652

Haber, J. R., Jacob, T., & Heath, A. C. (2005). Paternal alcoholism and offspring conduct disorder: Evidence for the 'common genes' hypothesis. *Twin Research and Human Genetics*, *8*(2), 120–131. https://doi.org/10.1375/twin.8.2.120

Hails, K. A., Shaw, D. S., Leve, L. D., Ganiban, J. M., Reiss, D., Natsuaki, M. N., & Neiderhiser, J. M. (2019). Interaction between adoptive mothers' and fathers' depressive symptoms in risk for children's emerging problem behavior. *Social Development*, *28*(3), 725–742. https://doi.org/10.1111/sode.12352

Hammerton, G., Heron, J., Mahedy, L., Maughan, B., Hickman, M., & Murray, J. (2018). Low resting heart rate, sensation seeking and the course of antisocial behaviour across adolescence and young adulthood. *Psychological Medicine*, *48*(13), 2194–2201. https://doi.org/10.1017/s0033291717003683

Hannigan, L. J., Eilertsen, E. M., Gjerde, L. C., Reichborn-Kjennerud, T., Eley, T. C., Rijsdijk, F. V., Ystrom, E., & McAdams, T. A. (2018). Maternal prenatal depressive symptoms and risk for early-life psychopathology in offspring: Genetic analyses in the Norwegian Mother and Child Birth Cohort Study. *The Lancet Psychiatry*, *5*(10), 808–815. https://doi.org/10.1016/S2215-0366(18)30225-6

Hänninen, T., & Soininen, H. (1997). Age-associated memory impairment: Normal aging or warning of dementia? *Drugs & Aging*, *11*(6), 480–489. https://doi.org/10.2165/00002512-199711060-00007

Harden, K. P., D'Onofrio, B. M., Van Hulle, C., Turkheimer, E., Rodgers, J. L., Waldman, I. D., & Lahey, B. B. (2009). Population density and youth antisocial

behavior. *Journal of Child Psychology and Psychiatry*, *50*(8), 999–1008. https://doi.org/10.1111/j.1469-7610.2009.02044.x

Harden, K. P., Lynch, S. K., Turkheimer, E., Emery, R. E., D'Onofrio, B. M., Slutske, W. S., Waldron, M. D., Heath, A. C., Statham, D. J., & Martin, N. G. (2007). A behavior genetic investigation of adolescent motherhood and offspring mental health problems. *Journal of Abnormal Psychology*, *116*(4), 667–683. https://doi.org/10.1037/0021-843x.116.4.667

Harden, K. P., Turkheimer, E., Emery, R. E., D'Onofrio, B. M., Slutske, W. S., Heath, A. C., & Martin, N. G. (2007). Marital conflict and conduct problems in Children of Twins. *Child Development*, *78*(1), 1–18. https://doi.org/10.1111/j.1467-8624.2007.00982.x

Harding, J. F. (2015). Increases in maternal education and low-income children's cognitive and behavioral outcomes. *Developmental Psychology*, *51*(5), 583–599. https://doi.org/10.1037/a0038920

Harold, G. T., Elam, K. K., Lewis, G., Rice, F., & Thapar, A. (2012). Interparental conflict, parent psychopathology, hostile parenting, and child antisocial behavior: Examining the role of maternal versus paternal influences using a novel genetically sensitive research design. *Development and Psychopathology*, *24*(4), 1283–1295. https://doi.org/10.1017/s0954579412000703

Harold, G. T., Leve, L. D., Elam, K. K., Thapar, A., Neiderhiser, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2013). The Nature of Nurture: Disentangling Passive Genotype–Environment Correlation From Family Relationship Influences on Children's Externalizing Problems. *Journal of Family Psychology*, *27*(1), 12–21. https://doi.org/10.1037/a0031190

Harold, G. T., Rice, F., Hay, D. F., Boivin, J., Van Den Bree, M., & Thapar, A. (2011). Familial transmission of depression and antisocial behavior symptoms: Disentangling the contribution of inherited and environmental factors and testing the mediating role of parenting. *Psychological Medicine*, *41*(6), 1175–1185. https://doi.org/10.1017/S0033291710001753

Harris, G. E., Wood, M., Ystrom, E., & Nordeng, H. (2018). Prenatal triptan exposure and neurodevelopmental outcomes in 5-year-old children: Follow-up from the Norwegian Mother and Child Cohort Study. *Paediatric and Perinatal Epidemiology*, *32*(3), 247–255. https://doi.org/10.1111/ppe.12461

Haskins, A. R. (2015). Paternal incarceration and child-reported behavioral functioning at age 9. *Social Science Research*, *52*, 18–33. https://doi.org/10.1016/j.ssresearch.2015.01.001

Hawkins, J. D., Herrenkohl, T. I., Farrington, D. P., Brewer, D., Catalano, R. F., Harachi, T. W., & Cothern, L. (2000). *Predictors of youth violence.* (179065). Office of

Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice.

Hay, D. F., Pawlby, S., Waters, C. S., Perra, O., & Sharp, D. (2010). Mothers' antenatal depression and their children's antisocial outcomes. *Child Development*, *81*(1), 149–165. https://doi.org/10.1111/j.1467-8624.2009.01386.x

Hazebroek, B. C. M. van, Wermink, H., Domburgh, L. van, Keijser, J. W. de, Hoeve, M., & Popma, A. (2019). Biosocial studies of antisocial behavior: A systematic review of interactions between peri/prenatal complications, psychophysiological parameters, and social risk factors. *Aggression and Violent Behavior*, *47*, 169–188. https://doi.org/10.1016/j.avb.2019.02.016

Heeks, M., Reed, S., Tafsiri, M., & Prince, S. (2018). *The economic and social costs of crime second edition*. Home Office Research Report. https://www. gov. uk/government/publications/the-economic-and-socialcosts-of-crime

Heim, C., Newport, D. J., Heit, S., Graham, Y. P., Wilcox, M., Bonsall, R., Miller, A. H., & Nemeroff, C. B. (2000). Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *JAMA*, *284*(5), 592–597. https://doi.org/10.1001/jama.284.5.592

Hemani, G., Bowden, J., & Davey Smith, G. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Human Molecular Genetics*, *27*(R2), R195–R208. https://doi.org/10.1093/hmg/ddy163

Hemani, G., Tilling, K., & Smith, G. D. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genetics*, *13*(11), e1007081. https://doi.org/10.1371/journal.pgen.1007081

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., Tan, V. Y., Yarmolinsky, J., Shihab, H. A., Timpson, N. J., Evans, D. M., Relton, C., Martin, R. M., Smith, G. D., Gaunt, T. R., & Haycock, P. C. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *eLife*, *7*, e34408. https://doi.org/10.7554/elife.34408

Herbst, C. M., & Tekin, E. (2016). The impact of child-care subsidies on child development: Evidence from geographic variation in the distance to social service agencies. *Journal of Policy Analysis and Management*, *35*(1), 94–116. https://doi.org/10.1002/pam.21860

Herle, M., Micali, N., Abdulkadir, M., Loos, R., Bryant-Waugh, R., Hübel, C., Bulik, C. M., & Stavola, B. L. D. (2020). Identifying typical trajectories in longitudinal data: Modelling strategies and interpretations. *European Journal of Epidemiology*, *35*(3), 205–222. https://doi.org/10.1007/s10654-020-00615-6

Hernán, M. A. (2004). A definition of causal effect for epidemiological research. *Journal of Epidemiology and Community Health*, *58*(4), 265–271. https://doi.org/10.1136/jech.2002.006361

Hernán, M. A., Hernández-Díaz, S., & Robins, J. M. (2004). A Structural Approach to Selection Bias. *Epidemiology*, *15*(5), 615–625. https://doi.org/10.1097/01.ede.0000135174.63482.43

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, *183*(8), 758–764. https://doi.org/10.1093/aje/kwv254

Hernán, M. A., & Robins, J. M. (2020). *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC.

Hernán, M. A., Sauer, B. C., Hernández-Díaz, S., Platt, R., & Shrier, I. (2016). Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of Clinical Epidemiology*, *79*, 70–75. https://doi.org/10.1016/j.jclinepi.2016.04.014

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hipwell, A. E., Murray, J., Xiong, S., Stepp, S. D., & Keenan, K. E. (2016). Effects of adolescent childbearing on maternal depression and problem behaviors: A prospective, population-based study using risk-set propensity scores. *PLOS ONE*, *11*(5), e0155641. https://doi.org/10.1371/journal.pone.0155641

Hipwell, A., Keenan, K., Kasza, K., Loeber, R., Stouthamer-Loeber, M., & Bean, T. (2008). Reciprocal Influences Between Girls' Conduct Problems and Depression, and Parental Punishment and Warmth: A Six year Prospective Analysis. *Journal of Abnormal Child Psychology*, *36*(5), 663–677. https://doi.org/10.1007/s10802-007-9206-4

Hou, J., Chen, Z., Natsuaki, M. N., Li, X., Yang, X., Zhang, J., & Zhang, J. (2013). A longitudinal investigation of the associations among parenting, deviant peer affiliation, and externalizing behaviors: A monozygotic twin differences design. *Twin Research and Human Genetics*, *16*(3), 698–706. https://doi.org/10.1017/thg.2013.24

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), 136–149. https://doi.org/10.1002/ab.20300

Humphrey, J. L., & Root, E. D. (2017). Spatio-temporal neighborhood impacts on internalizing and externalizing behaviors in U.S. elementary school children:

Effect modification by child and family socio-demographics. *Social Science & Medicine*, *180*, 52–61. https://doi.org/10.1016/j.socscimed.2017.03.014

Hwang, L.-D., Davies, N. M., Warrington, N. M., & Evans, D. M. (2021). Integrating Family-Based and Mendelian Randomization Designs. *Cold Spring Harbor Perspectives in Medicine*, *11*(3), a039503. https://doi.org/10.1101/cshperspect.a039503

Ichikawa, K., Fujiwara, T., & Kawachi, I. (2017). It takes a village: Fixed-effects analysis of neighborhood collective efficacy and children's development. *Journal of Epidemiology*, *27*(8), 368–372. https://doi.org/10.1016/j.je.2016.08.018

Ip, H. F., Laan, C. M. van der, Krapohl, E. M. L., Brikell, I., Sánchez-Mora, C., Nolte, I. M., Pourcain, B. S., 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*(1), 413. https://doi.org/10.1038/s41398-021-01480-x

Jackson, D. B. (2018). Does TV viewing during toddlerhood predict social difficulties and conduct problems? *Infant and Child Development*, *27*(4), e2086. https://doi.org/10.1002/icd.2086

Jackson, D., Viechtbauer, W., & van Aert, R. C. M. (2024). Multistep estimators of the between-study covariance matrix under the multivariate random-effects model for meta-analysis. *Statistics in Medicine*, *43*(4), 756–773. https://doi.org/10.1002/sim.9985

Jaffee, S. R., Hanscombe, K. B., Haworth, C. M. A., Davis, O. S. P., & Plomin, R. (2012). Chaotic Homes and Children's Disruptive Behavior. *Psychological Science*, *23*(6), 643–650. https://doi.org/10.1177/0956797611431693

Jaffee, S. R., Strait, L. B., & Odgers, C. L. (2012). From correlates to causes: Can quasi-experimental studies and statistical innovations bring us closer to identifying the causes of antisocial behavior? *Psychological Bulletin*, *138*(2), 272–295. https://doi.org/10.1037/a0026020

Jaffee, S. R., Van Hulle, C., & Rodgers, J. L. (2011). Effects of nonmaternal care in the first 3 years on children's academic skills and behavioral functioning in childhood and early adolescence: A sibling comparison study. *Child Development*, *82*(4), 1076–1091. https://doi.org/10.1111/j.1467-8624.2011.01611.x

Jandoc, R., Burden, A. M., Mamdani, M., Lévesque, L. E., & Cadarette, S. M. (2015). Interrupted time series analysis in drug utilization research is increasing:

Systematic review and recommendations. *Journal of Clinical Epidemiology*, *68*(8), 950–956. https://doi.org/10.1016/j.jclinepi.2014.12.018

Jennings, J. R., Pardini, D. A., & Matthews, K. A. (2017). Heart rate, health, and hurtful behavior: Heart rate, health, and hurtful behavior. *Psychophysiology*, *54*(3), 399–408. https://doi.org/10.1111/psyp.12802

Jennings, W. G., Piquero, A. R., & Farrington, D. P. (2013). Does resting heart rate at age 18 distinguish general and violent offending up to age 50? Findings from the Cambridge Study in Delinquent Development. *Journal of Criminal Justice*, *41*(4), 213–219. https://doi.org/10.1016/j.jcrimjus.2013.05.003

Jeong, J., Franchett, E. E., Ramos De Oliveira, C. V., Rehmani, K., & Yousafzai, A. K. (2021). Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLOS Medicine*, *18*(5), e1003602. https://doi.org/10.1371/journal.pmed.1003602

Jeong, J., McCoy, D. C., Yousafzai, A. K., Salhi, C., & Fink, G. (2016). Paternal stimulation and early child development in low- and middle-income countries. *Pediatrics*, *138*(4), e20161357. https://doi.org/10.1542/peds.2016-1357

Karwatowska, L., Frach, L., Schoeler, T., Tielbeek, J. J., Murray, J., de Geus, E., Viding, E., & Pingault, J.-B. (2023). Resting heart rate and antisocial behaviour: A Mendelian randomisation study. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-37123-y

Karwatowska, L., Russell, S., Solmi, F., Stavola, B. L. D., Jaffee, S., Pingault, J.-B., & Viding, E. (2020). Risk factors for disruptive behaviours: Protocol for a systematic review and meta-analysis of quasi-experimental evidence. *BMJ Open*, *10*(9), e038258. https://doi.org/10.1136/bmjopen-2020-038258

Karwatowska, L., Schoeler, T., Viding, E., & Pingault, J.-B. (2019). *Evaluating the causal effect of heart rate on antisocial behavior: A multivariable Mendelian randomization study*. *49*(6), 544–544.

Karwatowska, L., Solmi, F., Baldwin, J., Jaffee, S., Viding, E., Pingault, J.-B., & Stavola, B. L. D. (2023). P08 Positive and negative parenting and offspring disruptive behaviour: A meta-analysis of quasi-experimental evidence. *Journal of Epidemiology and Community Health*, *77*, A58–A58. https://doi.org/10.1136/jech-2023-SSMabstracts.117

Kavish, N., Bergstrøm, H., Piquero, A. R., Farrington, D. P., & Boutwell, B. B. (2020). The longitudinal association between resting heart rate and psychopathic traits from a normative personality perspective. *American Journal of Criminal Justice*, *45*(3), 410–425. https://doi.org/10.1007/s12103-019-09504-5

Kavish, N., Fu, Q. J., Vaughn, M. G., Qian, Z., & Boutwell, B. B. (2019). Resting heart rate and psychopathy: Findings from the Add Health Survey. *International*

Journal of Offender Therapy and Comparative Criminology, *63*(4), 543–557. https://doi.org/10.1177/0306624X18806748

Keil, V., & Price, J. M. (2009). Social information-processing patterns of maltreated children in two social domains. *Journal of Applied Developmental Psychology*, *30*(1), 43–52. https://doi.org/10.1016/j.appdev.2008.10.003

Kendler, K. S., Lönn, S. L., Sundquist, J., & Sundquist, K. (2021). The causal nature of the association between resting pulse in late adolescence and risk for internalizing and externalizing disorders: A co-relative analysis in a national male Swedish sample. *Psychological Medicine*, *51*(11), 1822–1828. https://doi.org/10.1017/s0033291720000549

Kendler, K. S., Ohlsson, H., Sundquist, J., & Sundquist, K. (2020). Maternal half-sibling families with discordant fathers: A contrastive design assessing cross-generational paternal genetic transmission of alcohol use disorder, drug abuse and major depression. *Psychological Medicine*, *50*(6), 973–980. https://doi.org/10.1017/S0033291719000874

Kerr, D. C. R., Leve, L. D., Harold, G. T., Natsuaki, M. N., Neiderhiser, J. M., Shaw, D. S., & Reiss, D. (2013). Influences of biological and adoptive mothers' depression and antisocial behavior on adoptees' early behavior trajectories. *Journal of Abnormal Child Psychology*, *41*(5), 723–734. https://doi.org/10.1007/s10802-013-9711-6

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., Aguilar-Gaxiola, S., Alhamzawi, A. O., Alonso, J., Angermeyer, M., Benjet, C., Bromet, E., Chatterji, S., Girolamo, G. de, Demyttenaere, K., Fayyad, J., Florescu, S., Gal, G., Gureje, O., … Williams, D. R. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *The British Journal of Psychiatry*, *197*(5), 378–385. https://doi.org/10.1192/bjp.bp.110.080499

Keyes, M., Legrand, L. N., Iacono, W. G., & McGue, M. (2008). Parental smoking and adolescent problem behavior: An adoption study of general and specific effects. *American Journal of Psychiatry*, *165*(10), 1338–1344. https://doi.org/10.1176/appi.ajp.2008.08010125

King, C. (2018). Food insecurity and child behavior problems in fragile families. *Economics & Human Biology*, *28*, 14–22. https://doi.org/10.1016/j.ehb.2017.11.002

King, S. M., Keyes, M., Malone, S. M., Elkins, I., Legrand, L. N., Iacono, W. G., & McGue, M. (2009). Parental alcohol dependence and the transmission of adolescent behavioral disinhibition: A study of adoptive and non-adoptive

families. *Addiction*, *104*(4), 578–586. https://doi.org/10.1111/j.1360-0443.2008.02469.x

Klahr, A. M., McGue, M., Iacono, W. G., & Burt, S. A. (2011). The association between parent-child conflict and adolescent conduct problems over time: Results from a longitudinal adoption study. *Journal of Abnormal Psychology*, *120*(1), 46–56. https://doi.org/10.1037/a0021350

Klahr, A. M., Rueter, M. A., McGue, M., Iacono, W. G., & Burt, S. A. (2011). The relationship between parent-child conflict and adolescent antisocial behavior: Confirming shared environmental mediation. *Journal of Abnormal Child Psychology*, *39*(5), 683–694. https://doi.org/10.1007/s10802-011-9505-7

Knudsen, A. K., Ystrom, E., Skogen, J. C., & Torgersen, L. (2015). Maternal heavy alcohol use and toddler behavior problems: A fixed effects regression analysis. *European Child & Adolescent Psychiatry*, *24*(10), 1269–1277. https://doi.org/10.1007/s00787-015-0677-5

Koss, K. J., & Gunnar, M. R. (2018). Annual Research Review: Early adversity, the HPA axis, and child psychopathology. *Journal of Child Psychology and Psychiatry*, *59*(4), 327–346. https://doi.org/10.1111/jcpp.12784

Krug, E. G., Mercy, J. A., Dahlberg, L. L., & Zwi, A. B. (2002). The world report on violence and health. *The Lancet*, *360*(9339), 1083–1088. https://doi.org/10.1016/S0140-6736(02)11133-0

Lacey, R. E., Howe, L. D., Kelly-Irving, M., Bartley, M., & Kelly, Y. (2022). The clustering of Adverse Childhood Experiences in the Avon Longitudinal Study of Parents and Children: Are gender and poverty important? *Journal of Interpersonal Violence*, *37*(5–6), 2218–2241. https://doi.org/10.1177/0886260520935096

Lacey, R. E., & Minnis, H. (2020). Practitioner Review: Twenty years of research with adverse childhood experience scores—Advantages, disadvantages and applications to practice. *Journal of Child Psychology and Psychiatry*, *61*(2), 116–130. https://doi.org/10.1111/jcpp.13135

Lai, E. T. C., Wickham, S., Law, C., Whitehead, M., Barr, B., & Taylor-Robinson, D. (2019). Poverty dynamics and health in late childhood in the UK: Evidence from the Millennium Cohort Study. *Archives of Disease in Childhood*, *104*(11), 1049–1055. https://doi.org/10.1136/archdischild-2018-316702

Lane, C., Hogg, E., Karwatowska, L. A., French, L., Ranieri, V. F., Jesnick, L. G. D., Roberts, C., Scott, S., Senior, R., Skinner, G. C., & Kennedy, E. M. M. (2023). Personalised interventions for subgroups of children with conduct problems. *The Cochrane Database of Systematic Reviews*, *4*(4), CD012746. https://doi.org/10.1002/14651858.CD012746.pub2

Lanier, P., Maguire-Jack, K., Lombardi, B., Frey, J., & Rose, R. A. (2018). Adverse childhood experiences and child health outcomes: Comparing cumulative risk and latent class approaches. *Maternal and Child Health Journal*, *22*(3), 288–297. https://doi.org/10.1007/s10995-017-2365-1

Lansford, J. E., Miller-Johnson, S., Berlin, L. J., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2007). Early physical abuse and later violent delinquency: A prospective longitudinal study. *Child Maltreatment*, *12*(3), 233–245. https://doi.org/10.1177/1077559507301841

Latham, R. M., Mark, K. M., & Oliver, B. R. (2017). A harsh parenting team? Maternal reports of coparenting and coercive parenting interact in association with children's disruptive behaviour. *Journal of Child Psychology and Psychiatry*, *58*(5), 603–611. https://doi.org/10.1111/jcpp.12665

Latvala, A., Kuja-Halkola, R., Almqvist, C., Larsson, H., & Lichtenstein, P. (2015). A longitudinal study of resting heart rate and violent criminality in more than 700000 men. *JAMA Psychiatry*, *72*(10), 971–978. https://doi.org/10.1001/jamapsychiatry.2015.1165

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Smith, G. D. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*, *27*(8), 1133–1163. https://doi.org/10.1002/sim.3034

Lawlor, D. A., Tilling, K., & Davey Smith, G. (2017). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, 1866–1886. https://doi.org/10.1093/ije/dyw314

Lawlor, D. A., Wade, K. H., Borges, M. C., Palmer, T., Hartwig, F. P., Hemani, G., & Bowden, J. (2019). A Mendelian Randomization dictionary: Useful definitions and descriptions for undertaking, understanding and interpreting Mendelian Randomization studies. *A Mendelian Randomization Dictionary*. https://doi.org/10.31219/osf.io/6yzs7

Lee, D., & McLanahan, S. (2015). Family structure transitions and child development: Instability, selection, and population heterogeneity. *American Sociological Review*, *80*(4), 738–763. https://doi.org/10.1177/0003122415592129

Lee, K. T. H., Lewis, R. W., Kataoka, S., Schenke, K., & Vandell, D. L. (2018). Out-of-school time and behaviors during adolescence. *Journal of Research on Adolescence*, *28*(2), 284–293. https://doi.org/10.1111/jora.12389

Leijten, P., Gardner, F., Melendez-Torres, G. J., Van Aar, J., Hutchings, J., Schulz, S., Knerr, W., & Overbeek, G. (2019). Meta-analyses: Key parenting program components for disruptive child behavior. *Journal of the American Academy of*

Child & Adolescent Psychiatry, *58*(2), 180–190. https://doi.org/10.1016/j.jaac.2018.07.900

Leijten, P., Melendez-Torres, G. J., & Gardner, F. (2022). Research Review: The most effective parenting program content for disruptive child behavior – a network meta-analysis. *Journal of Child Psychology and Psychiatry*, *63*(2), 132–142. https://doi.org/10.1111/jcpp.13483

Leve, L. D., Harold, G. T., Ge, X., Neiderhiser, J. M., Shaw, D., Scaramella, L. V., & Reiss, D. (2009). Structured parenting of toddlers at high versus low genetic risk: Two pathways to child problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(11), 1102–1109. https://doi.org/10.1097/CHI.0b013e3181b8bfc0

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, *126*(2), 309–337. https://doi.org/10.1037/0033-2909.126.2.309

Lewis, G. J., & Plomin, R. (2015). Heritable influences on behavioural problems from early childhood to mid-adolescence: Evidence for genetic stability and innovation. *Psychological Medicine*, *45*(10), 2171–2179. https://doi.org/10.1017/s0033291715000173

Lipscomb, S. T., Laurent, H., Neiderhiser, J. M., Shaw, D. S., Natsuaki, M. N., Reiss, D., & Leve, L. D. (2014). Genetic vulnerability interacts with parenting and early care and education to predict increasing externalizing behavior. *International Journal of Behavioral Development*, *38*(1), 70–80. https://doi.org/10.1177/0165025413508708

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-analysis*. Sage.

Liu, J., Raine, A., Wuerker, A., Venables, P. H., & Mednick, S. (2009). The association of birth complications and externalizing behavior in early adolescents: Direct and mediating effects. *Journal of Research on Adolescence : The Official Journal of the Society for Research on Adolescence*, *19*(1), 93–111. https://doi.org/10.1111/j.1532-7795.2009.00583.x

Lombardi, C. M., & Coley, R. L. (2014). Early maternal employment and children's school readiness in contemporary families. *Developmental Psychology*, *50*(8), 2071–2084. https://doi.org/10.1037/a0037106

Long, E. C., Aggen, S. H., Gardner, C., & Kendler, K. S. (2015). Differential parenting and risk for psychopathology: A monozygotic twin difference approach. *Social Psychiatry and Psychiatric Epidemiology*, *50*(10), 1569–1576. https://doi.org/10.1007/s00127-015-1065-7

Lorber, M. F. (2004). Psychophysiology of aggression, psychopathy, and conduct problems: A meta-analysis. *Psychological Bulletin*, *130*(4), 531–552. https://doi.org/10.1037/0033-2909.130.4.531

Lund, I. O., Moen Eilertsen, E., Gjerde, L. C., Røysamb, E., Wood, M., Reichborn-Kjennerud, T., & Ystrom, E. (2019). Is the association between maternal alcohol consumption in pregnancy and pre-school child behavioural and emotional problems causal? Multiple approaches for controlling unmeasured confounding. *Addiction*, *114*(6), 1004–1014. https://doi.org/10.1111/add.14573

Lundahl, B. W., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research on Social Work Practice*, *18*(2), 97–106. https://doi.org/10.1177/1049731507309828

Lysenko, L. J., Barker, E. D., & Jaffee, S. R. (2013). Sex differences in the relationship between harsh discipline and conduct problems. *Social Development*, *22*(1), 197–214. https://doi.org/10.1111/sode.12002

Ma, J., Grogan-Kaylor, A., & Lee, S. J. (2018). Associations of neighborhood disorganization and maternal spanking with children's aggression: A fixed-effects regression analysis. *Child Abuse & Neglect*, *76*, 106–116. https://doi.org/10.1016/j.chiabu.2017.10.013

Ma, J., Grogan-Kaylor, A., & Lee, S. J. (2020). Does community violence exposure moderate the associations between maternal spanking and early child behavior problems? *Aggressive Behavior*, *46*(3), 210–219. https://doi.org/10.1002/ab.21882

Maas, C., Herrenkohl, T. I., & Sousa, C. (2008). Review of research on child maltreatment and violence in youth. *Trauma, Violence, & Abuse*, *9*(1), 56–67. https://doi.org/10.1177/1524838007311105

MacKinnon, N., Kingsbury, M., Mahedy, L., Evans, J., & Colman, I. (2018). The association between prenatal stress and externalizing symptoms in childhood: Evidence from the Avon Longitudinal Study of Parents and Children. *Biological Psychiatry*, *83*(2), 100–108. https://doi.org/10.1016/j.biopsych.2017.07.010

Mankuta, D., Goldner, I., & Knafo, A. (2010). Intertwin birth weight differences and conduct problems in early childhood. *Archives of Pediatrics & Adolescent Medicine*, *457–61*(5), 457–461. https://doi.org/10.1001/archpediatrics.2010.63

Marceau, K., Hajal, N., Leve, L. D., Reiss, D., Shaw, D. S., Ganiban, J. M., Mayes, L. C., & Neiderhiser, J. M. (2013). Measurement and associations of pregnancy risk factors with genetic influences, postnatal environmental influences, and toddler behavior. *International Journal of Behavioral Development*, *37*(4), 366–375. https://doi.org/10.1177/0165025413489378

Mark, K. M., & Pike, A. (2017). Links between marital quality, the mother–child relationship and child behavior. *International Journal of Behavioral Development*, *41*(2), 285–294. https://doi.org/10.1177/0165025416635281

Marmot, M. (2020). Health equity in England: The Marmot review 10 years on. *BMJ*, *368*, m693. https://doi.org/10.1136/bmj.m693

Marmot, M., Friel, S., Bell, R., Houweling, T. A., & Taylor, S. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *The Lancet*, *372*(9650), 1661–1669. https://doi.org/10.1016/S0140-6736(08)61690-6

Mason, D. A., & Frick, P. J. (1994). The heritability of antisocial behavior: A meta-analysis of twin and adoption studies. *Journal of Psychopathology and Behavioral Assessment*, *16*(4), 301–323. https://doi.org/10.1007/BF02239409

Matthay, E. C., Hagan, E., Gottlieb, L. M., Tan, M. L., Vlahov, D., Adler, N. E., & Glymour, M. M. (2020). Alternative causal inference methods in population health research: Evaluating tradeoffs and triangulating evidence. *SSM - Population Health*, *10*, 100526. https://doi.org/10.1016/j.ssmph.2019.100526

Matthews, A. A., Danaei, G., Islam, N., & Kurth, T. (2022). Target trial emulation: Applying principles of randomised trials to observational studies. *BMJ*, *378*, e071108. https://doi.org/10.1136/bmj-2022-071108

Matthys, W., Vanderschuren, L. J. M. J., & Schutter, D. J. L. G. (2013). The neurobiology of oppositional defiant disorder and conduct disorder: Altered functioning in three mental domains. *Development and Psychopathology*, *25*(1), 193–207. https://doi.org/10.1017/s0954579412000272

Maughan, B., & Kim-Cohen, J. (2005). Continuities between childhood and adult life. *British Journal of Psychiatry*, *187*(4), 301–303. https://doi.org/10.1192/bjp.187.4.301

Maughan, B., Stafford, M., Shah, I., & Kuh, D. (2014). Adolescent conduct problems and premature mortality: Follow-up to age 65 years in a national birth cohort. *Psychological Medicine*, *44*(5), 1077–1086. https://doi.org/10.1017/S0033291713001402

McAdams, T. A., Rijsdijk, F. V., Neiderhiser, J. M., Narusyte, J., Shaw, D. S., Natsuaki, M. N., Spotts, E. L., Ganiban, J. M., Reiss, D., Leve, L. D., Lichtenstein, P., & Eley, T. C. (2015). The relationship between parental depressive symptoms and offspring psychopathology: Evidence from a children-of-twins study and an adoption study. *Psychological Medicine*, *45*(12), 2583–2594. https://doi.org/10.1017/S0033291715000501

McAdams, T. A., Rijsdijk, F. V., Zavos, H. M. S., & Pingault, J.-B. (2021). Twins and causal inference: Leveraging nature's experiment. *Cold Spring Harbor*

Perspectives in Medicine, *11*(6), a039552. https://doi.org/10.1101/cshperspect.a039552

McCrory, C., & Layte, R. (2012). Prenatal exposure to maternal smoking and childhood behavioural problems: A quasi-experimental approach. *Journal of Abnormal Child Psychology*, *40*(8), 1277–1288. https://doi.org/10.1007/s10802-012-9640-9

McCrory, E., Brito, S. A. D., & Viding, E. (2012). The link between child abuse and psychopathology: A review of neurobiological and genetic research. *Journal of the Royal Society of Medicine*, *105*(4), 151–156. https://doi.org/10.1258/jrsm.2011.110222

McCrory, E., Foulkes, L., & Viding, E. (2022). Social thinning and stress generation after childhood maltreatment: A neurocognitive social transactional model of psychiatric vulnerability. *The Lancet Psychiatry*, *9*(10), 828–837. https://doi.org/10.1016/S2215-0366(22)00202-4

McCrory, E., & Viding, E. (2015). The theory of latent vulnerability: Reconceptualizing the link between childhood maltreatment and psychiatric disorder. *Development and Psychopathology*, *27*(2), 493–505. https://doi.org/10.1017/S0954579415000115

McGuire, J., Kaiser, C., & Bach-Mortensen, A. M. (2022). A systematic review and meta-analysis of the impact of cash transfers on subjective well-being and mental health in low- and middle-income countries. *Nature Human Behaviour*, *6*(3), Article 3. https://doi.org/10.1038/s41562-021-01252-z

McLaughlin, K. A. (2016). Future Directions in Childhood Adversity and Youth Psychopathology. *Journal of Clinical Child & Adolescent Psychology*, *45*(3), 361–382. https://doi.org/10.1080/15374416.2015.1110823

McLaughlin, K. A., Colich, N. L., Rodman, A. M., & Weissman, D. G. (2020). Mechanisms linking childhood trauma exposure and psychopathology: A transdiagnostic model of risk and resilience. *BMC Medicine*, *18*(1), 96. https://doi.org/10.1186/s12916-020-01561-6

McLaughlin, K. A., Greif Green, J., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2012). Childhood adversities and first onset of psychiatric disorders in a national sample of US adolescents. *Archives of General Psychiatry*, *69*(11), 1151–1160. https://doi.org/10.1001/archgenpsychiatry.2011.2277

McLaughlin, K. A., & Lambert, H. K. (2017). Child trauma exposure and psychopathology: Mechanisms of risk and resilience. *Current Opinion in Psychology*, *14*, 29–34. https://doi.org/10.1016/j.copsyc.2016.10.004

Menting, A. T. A., Orobio De Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child

behavior: A meta-analytic review. *Clinical Psychology Review*, *33*(8), 901–913. https://doi.org/10.1016/j.cpr.2013.07.006

Meunier, J. C., Bisceglia, R., & Jenkins, J. M. (2012). Differential parenting and children's behavioral problems: Curvilinear associations and mother-father combined effects. *Developmental Psychology*, *48*(4), 987–1002. https://doi.org/10.1037/a0026321

Meyer, J. M., Rutter, M., Silberg, J. L., Maes, H. H., Simonoff, E., Shillady, L. L., Pickles, A., Hewitt, J. K., & Eaves, L. J. (2000). Familial aggregation for conduct disorder symptomatology: The role of genes, marital discord and family adaptability. *Psychological Medicine*, *30*(4), 759–774. https://doi.org/10.1017/S0033291799002408

Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M., & Jansen, W. (2012). Validity and Reliability of the Strengths and Difficulties Questionnaire in 5–6 Year Olds: Differences by Gender or by Parental Education? *PLoS ONE*, *7*(5), e36805. https://doi.org/10.1371/journal.pone.0036805

Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture on human aggression. *Journal of Personality and Social Psychology*, *72*(1), 207–217. https://doi.org/10.1037/0022-3514.72.1.207

Mingebach, T., Kamp-Becker, I., Christiansen, H., & Weber, L. (2018). Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PLoS ONE*, *13*(9), e0202855. https://doi.org/10.1371/journal.pone.0202855

Misheva, V., Webbink, D., & Martin, N. G. (2017). The effect of child maltreatment on illegal and problematic behaviour: New evidence on the 'cycle of violence' using twins data. *Journal of Population Economics*, *30*(4), 1035–1067. https://doi.org/10.1007/s00148-017-0642-3

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

Moffitt, T. E. (2005). Genetic and environmental influences on antisocial behaviors: Evidence from behavioral-genetic research. *Advances in Genetics*, *55*, 41–104. https://doi.org/10.1016/s0065-2660(05)55003-x

Moffitt, T. E., Arseneault, L., Jaffee, S. R., Kim-Cohen, J., Koenen, K. C., Odgers, C. L., Slutske, W. S., & Viding, E. (2008). Research Review: DSM-V conduct disorder: research needs for an evidence base. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(1), 3. https://doi.org/10.1111/j.1469-7610.2007.01823.x

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex Differences in Antisocial Behaviour: Conduct Disorder, Delinquency, and Violence in the Dunedin Longitudinal Study* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511490057

Møllegaard, S. (2020). The effect of birth weight on behavioral problems in early adolescence: New evidence from monozygotic twins. *Economics & Human Biology*, *36*, 100828. https://doi.org/10.1016/j.ehb.2019.100828

Monnet, J. (2019). The effect of preschool participation on intellectual and behavioral disorder diagnoses: Evidence from surveys on children's health. *Economics of Education Review*, *68*, 136–147. https://doi.org/10.1016/j.econedurev.2018.12.003

Morcillo, C., Duarte, C. S., Shen, S., Blanco, C., Canino, G., & Bird, H. R. (2011). Parental familism and antisocial behaviors: Development, gender, and potential mechanisms. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(5), 471–479. https://doi.org/10.1016/j.jaac.2011.01.014

Mors, O., Perto, G. P., & Mortensen, P. B. (2011). The Danish Psychiatric Central Research Register. *Scandinavian Journal of Public Health*, *39*(7_suppl), 54–57. https://doi.org/10.1177/1403494810395825

Mostafa, T., Gambaro, L., & Joshi, H. (2018). The impact of complex family structure on child well-being: Evidence from siblings. *Journal of Marriage and Family*, *80*(4), 902–918. https://doi.org/10.1111/jomf.12456

MQ Transforming Mental Health. (2021). *UK Mental Health Research Funding 2014–2017 [version 1; not peer reviewed].* AMRC Open Res. https://doi.org/10.21955/amrcopenres.1114943.1

Munafò, M. R., & Smith, G. D. (2018). Robust research needs many lines of evidence. *Nature*, *553*(7689), 399–401. https://doi.org/10.1038/d41586-018-01023-3

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. *International Journal of Epidemiology*, *47*(1), 226–235. https://doi.org/10.1093/ije/dyx206

Murray, J., Burgess, S., Zuccolo, L., Hickman, M., Gray, R., & Lewis, S. J. (2016). Moderate alcohol drinking in pregnancy increases risk for children's persistent conduct problems: Causal effects in a Mendelian randomisation study. *Journal of Child Psychology and Psychiatry*, *57*(5), 575–584. https://doi.org/10.1111/jcpp.12486

Murray, J., & Farrington, D. P. (2010). Risk factors for conduct disorder and delinquency: Key findings from longitudinal studies. *The Canadian Journal of Psychiatry*, *55*(10), 633–642. https://doi.org/10.1177/070674371005501003

Murray, J., Hallal, P. C., Mielke, G. I., Raine, A., Wehrmeister, F. C., Anselmi, L., & Barros, F. C. (2016). Low resting heart rate is associated with violence in late adolescence: A prospective birth cohort study in Brazil. *International Journal of Epidemiology*, *45*(2), 491–500. https://doi.org/10.1093/ije/dyv340

Murray, J., Shenderovich, Y., Gardner, F., Mikton, C., Derzon, J. H., Liu, J., & Eisner, M. (2018). Risk factors for antisocial behavior in low- and middle-income countries: A systematic review of longitudinal studies. *Crime Justice*, *47*(1), 255–364. https://doi.org/10.1086/696590

Nagin, D. S., Jones, B. L., Passos, V. L., & Tremblay, R. E. (2018). Group-based multi-trajectory modeling. *Statistical Methods in Medical Research*, *27*(7), 2015–2023. https://doi.org/10.1177/0962280216673085

Naimi, A. I., Cole, S. R., & Kennedy, E. H. (2016). An introduction to g methods. *International Journal of Epidemiology*, *46*(2), 756–762. https://doi.org/10.1093/ije/dyw323

Narusyte, J., Neiderhiser, J. M., Andershed, A.-K., D'Onofrio, B. M., Reiss, D., Spotts, E., Ganiban, J., & Lichtenstein, P. (2011). Parental criticism and externalizing behavior problems in adolescents– the role of environment and genotype-environment correlation. *Journal of Abnormal Psychology*, *120*(2), 365–376. https://doi.org/10.1037/a0021815

National Institute for Health and Care Excellence. (2013). *Antisocial behaviour and conduct disorders in children and young people: Recognition and management*.

Neiderhiser, J. M., Marceau, K., De Araujo-Greecher, M., Ganiban, J. M., Mayes, L. C., Shaw, D. S., Reiss, D., & Leve, L. D. (2016). Estimating the roles of genetic risk, perinatal risk, and marital hostility on early childhood adjustment: Medical records and self-reports. *Behavior Genetics*, *46*(3), 334–352. https://doi.org/10.1007/s10519-016-9788-0

Nelson, E. C., Heath, A. C., Madden, P. A. F., Cooper, M. L., Dinwiddie, S. H., Bucholz, K. K., Glowinski, A., McLaughlin, T., Dunne, M. P., Statham, D. J., & Martin, N. G. (2002). Association between self-reported childhood sexual abuse and adverse psychosocial outcomes: Results from a twin study. *Archives of General Psychiatry*, *59*(2), 139–145. https://doi.org/10.1001/archpsyc.59.2.139

Neugebauer, R., Hoek, H. W., & Susser, E. (1999). Prenatal exposure to wartime famine and development of antisocial personality disorder in early adulthood. *JAMA*, *Aug 4*(282(5)), 455–462. https://doi.org/10.1001/jama.282.5.455.

Newbury, J. B., Arseneault, L., Moffitt, T. E., Caspi, A., Danese, A., Baldwin, J. R., & Fisher, H. L. (2018). Measuring childhood maltreatment to predict early-adult psychopathology: Comparison of prospective informant-reports and

retrospective self-reports. *Journal of Psychiatric Research*, *96*, 57–64. https://doi.org/10.1016/j.jpsychires.2017.09.020

Nolte, I. M., Munoz, M. L., Tragante, V., Amare, A. T., Jansen, R., Vaez, A., Heyde, B. von der, Avery, C. L., Bis, J. C., Dierckx, B., Dongen, J. van, Gogarten, S. M., Goyette, P., Hernesniemi, J., Huikari, V., Hwang, S.-J., Jaju, D., Kerr, K. F., Kluttig, A., … Geus, E. J. C. de. (2017). Genetic loci associated with heart rate variability and their effects on cardiac disease risk. *Nature Communications*, *8*(1), 15805. https://doi.org/10.1038/ncomms15805

Nulman, I., Koren, G., Rovet, J., Barrera, M., Streiner, D. L., & Feldman, B. M. (2015). Neurodevelopment of children prenatally exposed to selective reuptake inhibitor antidepressants: Toronto sibling study. *The Journal of Clinical Psychiatry*, *76*(07), e842–e847. https://doi.org/10.4088/JCP.14m09240

Odgers, C. L., Caspi, A., Russell, M. A., Sampson, R. J., Arseneault, L., & Moffitt, T. E. (2012). Supportive parenting mediates neighborhood socioeconomic disparities in children's antisocial behavior from ages 5 to 12. *Development and Psychopathology*, *24*(3), 705–721. https://doi.org/10.1017/S0954579412000326

Oldenhof, H., Prätzlich, M., Ackermann, K., Baker, R., Batchelor, M., Baumann, S., Bernhard, A., Clanton, R., Dikeos, D., Dochnal, R., Fehlbaum, L. V., Fernández-Rivas, A., Geus, E. de, Gonzalez, K., Artaza-Lavesa, M. G. de, Guijarro, S., Gundlach, M., Herpertz-Dahlmann, B., Hervas, A., … Popma, A. (2019). Baseline autonomic nervous system activity in female children and adolescents with conduct disorder: Psychophysiological findings from the FemNAT-CD study. *Journal of Criminal Justice*, *65*, 1–10. https://doi.org/10.1016/j.jcrimjus.2018.05.011

Oliver, B. R. (2015). Unpacking externalising problems: Negative parenting associations for conduct problems and irritability. *BJPsych Open*, *1*(1), 42–47. https://doi.org/10.1192/bjpo.bp.115.000125

Oliver, B. R., Pike, A., & Plomin, R. (2008). Nonshared environmental influences on teacher-reported behaviour problems: Monozygotic twin differences in perceptions of the classroom. *Journal of Child Psychology and Psychiatry*, *49*(6), 646–653. https://doi.org/10.1111/j.1469-7610.2008.01891.x

Oliver, B. R., Trzaskowski, M., & Plomin, R. (2014). Genetics of parenting: The power of the dark side. *Developmental Psychology*, *50*(4), 1233–1240. https://doi.org/10.1037/a0035388

Orri, M., Tremblay, R. E., Japel, C., Boivin, M., Vitaro, F., Losier, T., Brendgen, M. R., Falissard, B., Melchior, M., & Côté, S. M. (2019). Early childhood child care and disruptive behavior problems during adolescence: A 17-year population-based

propensity score study. *Journal of Child Psychology and Psychiatry*, *60*(11), 1174–1182. https://doi.org/10.1111/jcpp.13065

Ortiz, J., & Raine, A. (2004). Heart Rate Level and Antisocial Behavior in Children and Adolescents: A Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(2), 154–162. https://doi.org/10.1097/00004583-200402000-00010

Ozenne, B. M. H., Scheike, T. H., Stærk, L., & Gerds, T. A. (2020). On the estimation of average treatment effects with right-censored time to event outcome and competing risks. *Biometrical Journal*, *62*(3), 751–763. https://doi.org/10.1002/bimj.201800298

Paine, A. L., Perra, O., Anthony, R., & Shelton, K. H. (2021). Charting the trajectories of adopted children's emotional and behavioral problems: The impact of early adversity and postadoptive parental warmth. *Development and Psychopathology*, *33*(3), 922–936. https://doi.org/10.1017/s0954579420000231

Palmer, R. H. C., Bidwell, L. C., Heath, A. C., Brick, L. A., Madden, P. A. F., & Knopik, V. S. (2016). Effects of maternal smoking during pregnancy on offspring externalizing problems: Contextual effects in a sample of female twins. *Behavior Genetics*, *46*(3), 403–415. https://doi.org/10.1007/s10519-016-9779-1

Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. F. (2014). Practitioner Review: Engaging fathers – recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry*, *55*(11), 1187–1212. https://doi.org/10.1111/jcpp.12280

Paradis, A. D., Shenassa, E. D., Papandonatos, G. D., Rogers, M. L., & Buka, S. L. (2017). Maternal smoking during pregnancy and offspring antisocial behaviour: Findings from a longitudinal investigation of discordant siblings. *Journal of Epidemiology and Community Health*, *71*(9), 889–896. https://doi.org/10.1136/jech-2016-208511

Pavlov, K. A., Chistiakov, D. A., & Chekhonin, V. P. (2012). Genetic determinants of aggression and impulsivity in humans. *Journal of Applied Genetics*, *53*(1), 61–82. https://doi.org/10.1007/s13353-011-0069-6

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *Vol. 82, No. 4*, 669–688.

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Pedersen, C. B. (2011). The Danish Civil Registration System. *Scandinavian Journal of Public Health*, *39*(7_suppl), 22–25. https://doi.org/10.1177/1403494810387965

Pemberton, C. K., Neiderhiser, J. M., Leve, L. D., Natsuaki, M. N., Shaw, D. S., Reiss, D., & Ge, X. (2010). Influence of parental depressive symptoms on adopted toddler behaviors: An emerging developmental cascade of genetic and environmental effects. *Development and Psychopathology*, *22*(4), 803–818. https://doi.org/10.1017/S0954579410000477

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(1), 51–87. https://doi.org/10.1111/j.1469-7610.1996.tb01380.x

Petersson, F., Baadsgaard, M., & Thygesen, L. C. (2011). Danish registers on personal labour market affiliation. *Scandinavian Journal of Public Health*, *39*(7_suppl), 95–98. https://doi.org/10.1177/1403494811408483

Pettit, G. S., Bates, J. E., & Dodge, K. A. (1997). Supportive parenting, ecological context, and children's adjustment: A seven-year longitudinal study. *Child Development*, *68*(5), 908. https://doi.org/10.2307/1132041

Pierce, B. L., & Burgess, S. (2013). Efficient design for Mendelian randomization studies: Subsample and 2-sample instrumental variable estimators. *American Journal of Epidemiology*, *178*(7), 1177–1184. https://doi.org/10.1093/aje/kwt084

Pike, A., McGuire, S., Hetherington, E. M., Reiss, D., & Plomin, R. (1996). Family environment and adolescent depressive symptoms and antisocial behavior: A multivariate genetic analysis. *Developmental Psychology*, *32*(4), 590–603. https://doi.org/10.1037/0012-1649.32.4.590

Pilling, S., Gould, N., Whittington, C., Taylor, C., & Scott, S. (2013). Recognition, intervention, and management of antisocial behaviour and conduct disorders in children and young people: Summary of NICE-SCIE guidance. *BMJ*, *346*, f1298. https://doi.org/10.1136/bmj.f1298

Pingault, J.-B., O'Reilly, P. F., Schoeler, T., Ploubidis, G. B., Rijsdijk, F., & Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*, *19*(9), 566–580. https://doi.org/10.1038/s41576-018-0020-3

Pingault, J.-B., Richmond, R., & Davey Smith, G. (2022). Causal inference with genetic data: Past, present, and future. *Cold Spring Harbor Perspectives in Medicine*, *12*(3), a041271. https://doi.org/10.1101/cshperspect.a041271

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

Piquero, A. R., Daigle, L. E., Gibson, C., Piquero, N. L., & Tibbetts, S. G. (2007). Are life-course-persistent offenders at risk for adverse health outcomes? *Journal of*

Research in Crime and Delinquency, *44*(2), 185–207. https://doi.org/10.1177/0022427806297739

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*(2), 309–322. https://doi.org/10.1037/0033-2909.84.2.309

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, *56*(3), 345–365. https://doi.org/10.1111/jcpp.12381

Polderman, T. J. C., Benyamin, B., Leeuw, C. A. de, Sullivan, P. F., Bochoven, A. van, 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

Polier, G. G., Vloet, T. D., Herpertz-Dahlmann, B., Laurens, K. R., & Hodgins, S. (2012). Comorbidity of conduct disorder symptoms and internalising problems in children: Investigating a community and a clinical sample. *European Child & Adolescent Psychiatry*, *21*(1), 31–38. https://doi.org/10.1007/s00787-011-0229-6

Portnoy, J., & Farrington, D. P. (2015). Resting heart rate and antisocial behavior: An updated systematic review and meta-analysis. *Aggression and Violent Behavior*, *22*, 33–45. https://doi.org/10.1016/j.avb.2015.02.004

Portnoy, J., Raine, A., Chen, F. R., Pardini, D., Loeber, R., & Jennings, J. R. (2014). Heart rate and antisocial behavior: The mediating role of impulse sensation seeking. *Criminology*, *52*(2), 292–311. https://doi.org/10.1111/1745-9125.12038

Powers, C. J., Bierman, K. L., & Coffman, D. L. (2016). Restrictive educational placements increase adolescent risks for students with early-starting conduct problems. *Journal of Child Psychology and Psychiatry*, *57*(8), 899–908. https://doi.org/10.1111/jcpp.12487

Prätzlich, M., Oldenhof, H., Steppan, M., Ackermann, K., Baker, R., Batchelor, M., Baumann, S., Bernhard, A., Clanton, R., Dikeos, D., Dochnal, R., Fehlbaum, L. V., Fernández-Rivas, A., Artaza-Lavesa, M. G. de, Gonzalez-Madruga, K., Guijarro, S., Gundlach, M., Herpertz-Dahlmann, B., Hervas, A., … Stadler, C. (2019). Resting autonomic nervous system activity is unrelated to antisocial

behaviour dimensions in adolescents: Cross-sectional findings from a European multi-centre study. *Journal of Criminal Justice*, *65*, 101536. https://doi.org/10.1016/j.jcrimjus.2018.01.004

Putnam, K. T., Harris, W. W., & Putnam, F. W. (2013). Synergistic childhood adversities and complex adult psychopathology. *Journal of Traumatic Stress*, *26*(4), 435–442. https://doi.org/10.1002/jts.21833

Raine, A. (1993). Crime and the Nature of Psychopathology. In *The Psychopathology of Crime* (pp. 1–26). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780080571485500058

Raine, A. (2002). Biosocial studies of antisocial and violent behavior in children and adults: A review. *Journal of Abnormal Child Psychology*, *30*(4), 311–326. https://doi.org/10.1023/a:1015754122318

Raine, A., Venables, P. H., & Mednick, S. A. (1997). Low Resting Heart Rate at Age 3 Years Predisposes to Aggression at Age 11 Years: Evidence From the Mauritius Child Health Project. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(10), 1457–1464. https://doi.org/10.1097/00004583-199710000-00029

Ramanathan, S., Balasubramanian, N., & Faraone, S. V. (2017). Familial transient financial difficulties during infancy and long-term developmental concerns. *Psychological Medicine*, *47*(12), 2197–2204. https://doi.org/10.1017/S0033291717000666

Ramiro, L. S., Madrid, B. J., & Brown, D. W. (2010). Adverse childhood experiences (ACE) and health-risk behaviors among adults in a developing country setting. *Child Abuse & Neglect*, *34*(11), 842–855. https://doi.org/10.1016/j.chiabu.2010.02.012

Reeves, B. C., Wells, G. A., & Waddington, H. (2017). Quasi-experimental study designs series—Paper 5: A checklist for classifying studies evaluating the effects on health interventions—A taxonomy without labels. *Journal of Clinical Epidemiology*, *89*, 30–42. https://doi.org/10.1016/j.jclinepi.2017.02.016

Reuben, A., Moffitt, T. E., Caspi, A., Belsky, D. W., Harrington, H., Schroeder, F., Hogan, S., Ramrakha, S., Poulton, R., & Danese, A. (2016). Lest we forget: Comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *Journal of Child Psychology and Psychiatry*, *57*(10), 1103–1112. https://doi.org/10.1111/jcpp.12621

Reuben, J. D., Shaw, D. S., Neiderhiser, J. M., Natsuaki, M. N., Reiss, D., & Leve, L. D. (2016a). Warm Parenting and Effortful Control in Toddlerhood: Independent and Interactive Predictors of School-Age Externalizing Behavior. *Journal of*

Abnormal Child Psychology, *44*(6), 1083–1096. https://doi.org/10.1007/s10802-015-0096-6

Reuben, J. D., Shaw, D. S., Neiderhiser, J. M., Natsuaki, M. N., Reiss, D., & Leve, L. D. (2016b). Warm parenting and effortful control in toddlerhood: Independent and interactive predictors of school-age externalizing behavior. *Journal of Abnormal Child Psychology*, *44*(6), 1083–1096. https://doi.org/10.1007/s10802-015-0096-6

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. https://doi.org/10.1037/0033-2909.128.3.490

Rice, F., Harold, G. T., Boivin, J., Van Den Bree, M., Hay, D. F., & Thapar, A. (2010). The links between prenatal stress and offspring development and psychopathology: Disentangling environmental and inherited influences. *Psychological Medicine*, *40*(2), 335–345. https://doi.org/10.1017/S0033291709005911

Richiardi, L., Pizzi, C., & Pearce, N. (2013). Commentary: Representativeness is usually not necessary and often should be avoided. *International Journal of Epidemiology*, *42*(4), 1018–1022. https://doi.org/10.1093/ije/dyt103

Richmond, M. K., & Stocker, C. M. (2003). Siblings' differential experiences of marital conflict and differences in psychological adjustment. *Journal of Family Psychology*, *17*(3), 339–350. https://doi.org/10.1037/0893-3200.17.3.339

Richmond, M. K., & Stocker, C. M. (2006). Associations between family cohesion and adolescent siblings' externalizing behavior. *Journal of Family Psychology*, *20*(4), 663–669. https://doi.org/10.1037/0893-3200.20.4.663

Richmond, M. K., & Stocker, C. M. (2009). Associations between siblings' differential family experiences and differences in psychological adjustment. *International Journal of Developmental Science*, *3*(2), 98–114. https://doi.org/10.3233/DEV-2009-3202

Riggins-Caspers, K. M., Cadoret, R. J., Knutson, J. F., & Langbehn, D. (2003). Biology-Environment Interaction and Evocative Biology-Environment Correlation: Contributions of Harsh Discipline and Parental Psychopathology to Problem Adolescent Behaviors. *Behavior Genetics*, *33*(3), 205–220. https://doi.org/10.1023/a:1023434206261

Rivenbark, J. G., Odgers, C. L., Caspi, A., Harrington, H., Hogan, S., Houts, R. M., Poulton, R., & Moffitt, T. E. (2018). The high societal costs of childhood conduct problems: Evidence from administrative records up to age 38 in a

longitudinal birth cohort. *Journal of Child Psychology and Psychiatry*, *59*(6), 703–710. https://doi.org/10.1111/jcpp.12850

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect. *Mathematical Modelling*, *7*(9), 1393–1512. https://doi.org/10.1016/0270-0255(86)90088-6

Rod, N. H., Bengtsson, J., Budtz-Jørgensen, E., Clipet-Jensen, C., Taylor-Robinson, D., Andersen, A.-M. N., Dich, N., & Rieckmann, A. (2020). Trajectories of childhood adversity and mortality in early adulthood: A population-based cohort study. *The Lancet*, *396*(10249), 489–497. https://doi.org/10.1016/s0140-6736(20)30621-8

Rod, N. H., Lange, T., & Petersen, A. H. (2021). Do sibling comparisons answer the causal question? In response to 'No causal associations between childhood family income and subsequent psychiatric disorders, substance misuse and violent crime arrests'. *International Journal of Epidemiology*, *51*(6), 2025–2026. https://doi.org/10.1093/ije/dyab235

Rodgers, M. A., & Pustejovsky, J. E. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*, *26*(2), 141–160. https://doi.org/10.1037/met0000300

Rodrigues, M., Sokolovic, N., Madigan, S., Luo, Y., Silva, V., Misra, S., & Jenkins, J. (2021). Paternal sensitivity and children's cognitive and socioemotional outcomes: A meta-analytic review. *Child Development*, *92*(2), 554–577. https://doi.org/10.1111/cdev.13545

Rolon-Arroyo, B., Arnold, D. H., Breaux, R. P., & Harvey, E. A. (2018). Reciprocal Relations Between Parenting Behaviors and Conduct Disorder Symptoms in Preschool Children. *Child Psychiatry & Human Development*, *49*(5), 786–799. https://doi.org/10.1007/s10578-018-0794-8

Roos, L. E., Fisher, P. A., Shaw, D. S., Kim, H. K., Neiderhiser, J. M., Reiss, D., Natsuaki, M. N., & Leve, L. D. (2016). Inherited and environmental influences on a childhood co-occurring symptom phenotype: Evidence from an adoption study. *Development and Psychopathology*, *28*(1), 111–125. https://doi.org/10.1017/S0954579415000322

Rothbaum, F., & Weisz, J. R. (1994). Parental caregiving and child externalizing behavior in nonclinical samples: A meta-analysis. *Psychological Bulletin*, *116*(1), 55–74. https://doi.org/10.1037/0033-2909.116.1.55

Rothman, K. J., Gallacher, J. E., & Hatch, E. E. (2013). Why representativeness should be avoided. *International Journal of Epidemiology*, *42*(4), 1012–1014. https://doi.org/10.1093/ije/dys223

Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine*, *26*(1), 20–36. https://doi.org/10.1002/sim.2739

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, *2*(3), 808–840. https://doi.org/10.1214/08-AOAS187

Rudolph, K. E., Shev, A., Paksarian, D., Merikangas, K. R., Mennitt, D. J., James, P., & Casey, J. A. (2019). Environmental noise and sleep and mental health outcomes in a nationally representative sample of urban US adolescents. *Environmental Epidemiology*, *3*(4), e056. https://doi.org/10.1097/EE9.0000000000000056

Rutter, M. (2007). *Identifying the Environmental Causes of Disease: How Should We Decide what to Believe and when to Take Action? Report Synopsis.*

Rutter, M. (2012). 'Natural Experiments' as a Means of Testing Causal Inferences. In *Causality* (pp. 253–272). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119945710.ch18

Sackett, D. L. (1979). Bias in analytic research. *Journal of Chronic Diseases*, *32*(1–2), 51–63. https://doi.org/10.1016/0021-9681(79)90012-2

Salvatore, J. E., & Dick, D. M. (2018). Genetic influences on conduct disorder. *Neuroscience & Biobehavioral Reviews*, *91*, 91–101. https://doi.org/10.1016/j.neubiorev.2016.06.034

Samek, D. R., Keyes, M. A., Hicks, B. M., Bailey, J., McGue, M., & Iacono, W. G. (2014). General and specific predictors of nicotine and alcohol dependence in early adulthood: Genetic and environmental influences. *Journal of Studies on Alcohol and Drugs*, *75*(4), 623–634. https://doi.org/10.15288/jsad.2014.75.623

Sampson, R. J., & Winter, A. S. (2018). Poisoned development: Assessing childhood lead exposure as a cause of crime in a birth cohort followed through adolescence: lead poisoning and crime. *Criminology*, *56*(2), 269–301. https://doi.org/10.1111/1745-9125.12171

Sanderson, E., Glymour, M. M., Holmes, M. V., Kang, H., Morrison, J., Munafò, M. R., Palmer, T., Schooling, C. M., Wallace, C., Zhao, Q., & Smith, G. D. (2022). Mendelian randomization. *Nature Reviews Methods Primers*, *2*(1), 6. https://doi.org/10.1038/s43586-021-00092-5

Sanderson, E., Richardson, T. G., Morris, T. T., Tilling, K., & Smith, G. D. (2022). Estimation of causal effects of a time-varying exposure at multiple time points through Multivariable Mendelian randomization. *PLoS Genetics*, *18(7)*, e1010290. https://doi.org/10.1371/journal.pgen.1010290

Schermerhorn, A. C., D'Onofrio, B. M., Turkheimer, E., Ganiban, J. M., Spotts, E. L., Lichtenstein, P., Reiss, D., & Neiderhiser, J. M. (2011). A genetically informed study of associations between family functioning and child psychosocial

adjustment. *Developmental Psychology*, *47*(3), 707–725. https://doi.org/10.1037/a0021362

Scherrer, J. F., Lin, N., Eisen, S. A., Goldberg, J., True, W. R., Lyons, M. J., & Tsuang, M. T. (1996). The association of antisocial personality symptoms with marijuana abuse/dependence. A monozygotic co-twin control study. *The Journal of Nervous and Mental Disease*, *184*(10), 611–615. https://doi.org/10.1097/00005053-199610000-00005

Schmidt, M., Schmidt, S. A. J., Sandegaard, J. L., Ehrenstein, V., Pedersen, L., & Sørensen, H. T. (2015). The Danish National Patient Registry: A review of content, data quality, and research potential. *Clinical Epidemiology*, *7*, 449–490. https://doi.org/10.2147/clep.s91125

Schoeler, T., Duncan, L., Cecil, C. M., Ploubidis, G. B., & Pingault, J.-B. (2018). Quasi-experimental evidence on short- and long-term consequences of bullying victimization: A meta-analysis. *Psychological Bulletin*, *144*(12), 1229–1246. https://doi.org/10.1037/bul0000171

Schoenfeld, D. (1982). Partial Residuals for The Proportional Hazards Regression Model. *Biometrika*, *69*(1), 239–241. https://doi.org/10.2307/2335876

Schoorl, J., Rijn, S. V., Wied, M. D., Goozen, S. H. M. V., & Swaab, H. (2016). Variability in emotional/behavioral problems in boys with oppositional defiant disorder or conduct disorder: The role of arousal. *European Child & Adolescent Psychiatry*, *25*(8), 821–830. https://doi.org/10.1007/s00787-015-0790-5

Scott, J. K., Nelson, J. A., & Dix, T. (2018). Interdependence among mothers, fathers, and children from early to middle childhood: Parents' sensitivity and children's externalizing behavior. *Developmental Psychology*, *54*(8), 1528–1541. https://doi.org/10.1037/dev0000525

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & Group, P.-P. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*, g7647. https://doi.org/10.1136/bmj.g7647

Shanahan, L., Copeland, W. E., Costello, E. J., & Angold, A. (2011). Child-, Adolescent-, and Young Adult-Onset Depressions: Differential Risk Factors in Development? *Psychological Medicine*, *41*(11), 2265–2274. https://doi.org/10.1017/S0033291711000675

Shelton, K. H., Collishaw, S., Rice, F. J., Harold, G. T., & Thapar, A. (2011). Using a genetically informative design to examine the relationship between breastfeeding and childhood conduct problems. *European Child & Adolescent Psychiatry*, *20*(11–12), 571–579. https://doi.org/10.1007/s00787-011-0224-y

Shelton, K. H., Harold, G. T., Fowler, T. A., Rice, F. J., Neale, M. C., Thapar, A., & Bree, M. B. M. van den. (2008). Parent–Child Relations, Conduct Problems and Cigarette Use in Adolescence: Examining the Role of Genetic and Environmental Factors on Patterns of Behavior. *Journal of Youth and Adolescence*, *37*(10), 1216–1228. https://doi.org/10.1007/s10964-007-9254-7

Shelton, K. H., Harold, G. T., Fowler, T. A., Rice, F. J., Neale, M. C., Thapar, A., & Van Den Bree, M. B. M. (2008). Parent-child relations, conduct problems and cigarette use in adolescence: Examining the role of genetic and environmental factors on patterns of behavior. *Journal of Youth and Adolescence*, *37*(10), 1216–1228. https://doi.org/10.1007/s10964-007-9254-7

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

Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., Garner, A. S., McGuinn, L., Pascoe, J., & Wood, D. L. (2012). The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *Pediatrics*, *129*(1), e232–e246. https://doi.org/10.1542/peds.2011-2663

Sijtsema, J. J., Veenstra, R., Lindenberg, S., Roon, A. M. van, Verhulst, F. C., Ormel, J., & Riese, H. (2010). Mediation of sensation seeking and behavioral inhibition on the relationship between heart rate and antisocial behavior: The TRAILS study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(5), 493–502. https://doi.org/10.1016/j.jaac.2010.02.005

Silwal, A. R., Engilbertsdottir, S., Cuesta, J., Newhouse, D., & Stewart, D. (2020). *Global estimate of children in monetary poverty: An update*. World Bank. http://documents.worldbank.org/curated/en/966791603123453576/Global-Estimate-of-Children-in-Monetary-Poverty-An-Update

Simonoff, E., Elander, J., Holmshaw, J., Pickles, A., Murray, R., & Rutter, M. (2004). Predictors of antisocial personality: Continuities from childhood to adult life. *The British Journal of Psychiatry*, *184*(2), 118–127. https://doi.org/10.1192/bjp.184.2.118

Singh, A. L., D'Onofrio, B. M., Slutske, W. S., Turkheimer, E., Emery, R. E., Harden, K. P., Heath, A. C., Madden, P. A. F., Statham, D. J., & Martin, N. G. (2011). Parental depression and offspring psychopathology: A Children of Twins study. *Psychological Medicine*, *41*(7), 1385–1395. https://doi.org/10.1017/S0033291710002059

Singham, T., Viding, E., Schoeler, T., Arseneault, L., Ronald, A., Cecil, C. M., McCrory, E., Rijsdijk, F., & Pingault, J.-B. (2017). Concurrent and longitudinal contribution of exposure to bullying in childhood to mental health: The role of vulnerability and resilience. *JAMA Psychiatry*, *74*(11), 1112–1119. https://doi.org/10.1001/jamapsychiatry.2017.2678

Sjölander, A., Frisell, T., & Öberg, S. (2022). Sibling Comparison Studies. *Annual Review of Statistics and Its Application*, *9*(1), 71–94. https://doi.org/10.1146/annurev-statistics-040120-024521

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, *57*(4), 193–216. https://doi.org/10.1037/h0054367

Skopp, N. A., McDonald, R., Manke, B., & Jouriles, E. N. (2005). Siblings in domestically violent families: Experiences of interparent conflict and adjustment problems. *Journal of Family Psychology*, *19*(2), 324–333. https://doi.org/10.1037/0893-3200.19.2.324

Skrivankova, V. W., Richmond, R. C., Woolf, B. A. R., Davies, N. M., Swanson, S. A., VanderWeele, T. J., Timpson, N. J., Higgins, J. P. T., Dimou, N., Langenberg, C., Loder, E. W., Golub, R. M., Egger, M., Smith, G. D., & Richards, J. B. (2021). Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): Explanation and elaboration. *BMJ*, *375*, n2233. https://doi.org/10.1136/bmj.n2233

Sonuga-Barke, E. J. S., Kennedy, M., Kumsta, R., Knights, N., Golm, D., Rutter, M., Maughan, B., Schlotz, W., & Kreppner, J. (2017). Child-to-adult neurodevelopmental and mental health trajectories after early life deprivation: The young adult follow-up of the longitudinal English and Romanian Adoptees study. *The Lancet*, *389*(10078), 1539–1548. https://doi.org/10.1016/S0140-6736(17)30045-4

Staff, J., Maggs, J. L., Bucci, R., & Mongilio, J. (2019). Changes in externalizing behaviors after children first have an alcoholic drink and first drink heavily. *Journal of Studies on Alcohol and Drugs*, *80*(4), 472–479. https://doi.org/10.15288/jsad.2019.80.472

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., Moher, D., Becker, B. J., Sipe, T. A., & Thacker, S. B. (2000). Meta-analysis of Observational Studies in Epidemiology: A Proposal for Reporting. *JAMA*, *283*(15), 2008–2012. https://doi.org/10.1001/jama.283.15.2008

Sundbakk, L. M., Wood, M., Gran, J. M., & Nordeng, H. (2019). Impact of prenatal exposure to benzodiazepines and z-hypnotics on behavioral problems at 5 years of age: A study from the Norwegian Mother and Child Cohort Study. *PLoS ONE*, *14*(6), e0217830. https://doi.org/10.1371/journal.pone.0217830

Taraban, L., Shaw, D. S., Leve, L. D., Natsuaki, M. N., Ganiban, J. M., Reiss, D., & Neiderhiser, J. M. (2019). Parental depression, overreactive parenting, and early childhood externalizing problems: Moderation by social support. *Child Development*, *90*(4), e468–e485. https://doi.org/10.1111/cdev.13027

Taylor-Robinson, D. C., Straatmann, V. S., & Whitehead, M. (2018). Adverse childhood experiences or adverse childhood socioeconomic conditions? *The Lancet Public Health*, *3*(6), e262–e263. https://doi.org/10.1016/s2468-2667(18)30094-x

Thapar, A., & Rice, F. (2021). Family-based designs that disentangle inherited factors from pre- and postnatal environmental exposures: In vitro fertilization, discordant sibling pairs, maternal versus paternal comparisons, and adoption designs. *Cold Spring Harbor Perspectives in Medicine*, *11*(3), a038877. https://doi.org/10.1101/cshperspect.a038877

The Lancet Public Health. (2021). Childhood adversity: A profound determinant of health. *The Lancet Public Health*, *6*(11), e780. https://doi.org/10.1016/S2468-2667(21)00237-1

Therneau, T. (2020). *A package for survival analysis in R* [Computer software]. Citeseer. https://CRAN.R-project.org/package=survival

Thomas, J., Brunton, J., & Graziosi, S. (2010). *EPPI-Reviewer 4.0: Software for research synthesis* [Computer software].

Thomson, R. M., Igelström, E., Purba, A. K., Shimonovich, M., Thomson, H., McCartney, G., Reeves, A., Leyland, A., Pearce, A., & Katikireddi, S. V. (2022). How do income changes impact on mental health and wellbeing for working-age adults? A systematic review and meta-analysis. *The Lancet Public Health*, *7*(6), e515–e528. https://doi.org/10.1016/S2468-2667(22)00058-5

Thornberry, T. P., Lizotte, A. J., Krohn, M. D., Farnworth, M., & Jang, S. J. (1994). Delinquent Peers, Beliefs, and Delinquent Behavior: A Longitudinal Test of Interactional Theory. *Criminology*, *32*(1), 47–83. https://doi.org/10.1111/j.1745-9125.1994.tb01146.x

Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: A meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry*, *57*(5), 585–595. https://doi.org/10.1111/jcpp.12499

Tielbeek, J. J., Uffelmann, E., Williams, B. S., Colodro-Conde, L., Gagnon, É., Mallard, T. T., Levitt, B., Jansen, P. R., Johansson, A., Sallis, H., Pistis, G., Saunders, G. R. B., Allegrini, A. G., Rimfeld, K., Konte, B., Klein, M., Hartmann, A. M., Salvatore, J. E., Nolte, I. M., … Posthuma, D. (2022). Uncovering the genetic architecture of broad antisocial behavior through a genome-wide association

study meta-analysis. *Molecular Psychiatry*, *27(11)*, 4453–4463. https://doi.org/10.1038/s41380-022-01793-3

Tore, E. C., Antoniou, E. E., Reed, K., Southwood, T. R., Smits, L., McCleery, J. P., & Zeegers, M. P. (2018). The association of intrapair birth-weight differences with internalizing and externalizing behavior problems. *Twin Research and Human Genetics*, *21*(3), 253–262. https://doi.org/10.1017/thg.2018.13

Travaglia, A., Bisaz, R., Sweet, E. S., Blitzer, R. D., & Alberini, C. M. (2016). Infantile amnesia reflects a developmental critical period for hippocampal learning. *Nature Neuroscience*, *19*(9), Article 9. https://doi.org/10.1038/nn.4348

Trønnes, J. N., Wood, M., Lupattelli, A., Ystrom, E., & Nordeng, H. (2020). Prenatal paracetamol exposure and neurodevelopmental outcomes in preschool-aged children. *Paediatric and Perinatal Epidemiology*, *34*(3), 247–256. https://doi.org/10.1111/ppe.12568

Ttofi, M. M., Farrington, D. P., Piquero, A. R., Lösel, F., DeLisi, M., & Murray, J. (2016). Intelligence as a protective factor against offending: A meta-analytic review of prospective longitudinal studies. *Journal of Criminal Justice*, *45*, 4–18. https://doi.org/10.1016/j.jcrimjus.2016.02.003

Tully, E. C., Iacono, W. G., & McGue, M. (2008). An adoption study of parental depression as an environmental liability for adolescent depression and childhood disruptive disorders. *American Journal of Psychiatry*, *165*(9), 1148–1154. https://doi.org/10.1176/appi.ajp.2008.07091438

Turney, K. (2012). Pathways of disadvantage: Explaining the relationship between maternal depression and children's problem behaviors. *Social Science Research*, *41*(6), 1546–1564. https://doi.org/10.1016/j.ssresearch.2012.06.003

Turney, K. (2017). The Unequal Consequences of Mass Incarceration for Children. *Demography*, *54*(1), 361–389. https://doi.org/10.1007/s13524-016-0543-1

VanderWeele, T. J. (2009). Concerning the Consistency Assumption in Causal Inference. *Epidemiology*, *20*(6), 880. https://doi.org/10.1097/EDE.0b013e3181bd5638

VanderWeele, T. J. (2019). Principles of confounder selection. *European Journal of Epidemiology*, *34*(3), 211–219. https://doi.org/10.1007/s10654-019-00494-6

VanderWeele, T. J. (2022). Constructed Measures and Causal Inference. *Epidemiology (Cambridge, Mass.)*, *33*(1), 141–151. https://doi.org/10.1097/EDE.0000000000001434

Verbanck, M., Chen, C.-Y., Neale, B., & Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature Genetics*, *50*(5), Article 5. https://doi.org/10.1038/s41588-018-0099-7

Viding, E., Fontaine, N. M. G., Oliver, B. R., & Plomin, R. (2009). Negative parental discipline, conduct problems and callous–unemotional traits: Monozygotic twin differences study. *British Journal of Psychiatry*, *195*(5), 414–419. https://doi.org/10.1192/bjp.bp.108.061192

Viding, E., & McCrory, E. (2020). Disruptive behavior disorders: The challenge of delineating mechanisms in the face of heterogeneity. *American Journal of Psychiatry*, *177*(9), 811–817. https://doi.org/10.1176/appi.ajp.2020.20070998

Viding, E., McCrory, E., Baskin-Sommers, A., Brito, S. D., & Frick, P. (2023). An 'embedded brain' approach to understanding antisocial behaviour. *Trends in Cognitive Sciences*, *28*(2), 159–171. https://doi.org/10.1016/j.tics.2023.08.013

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

Viechtbauer, W., & Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Wachs, T. D., Georgieff, M., Cusick, S., & McEwen, B. S. (2014). Issues in the timing of integrated early interventions: Contributions from nutrition, neuroscience, and psychological research. *Annals of the New York Academy of Sciences*, *1308*(1), 89–106. https://doi.org/10.1111/nyas.12314

Wadsworth, M. E. J. (1976). Delinquency, pulse rates and early emotional deprivation. *The British Journal of Criminology*, *16*(3), 245–256. https://doi.org/10.1093/oxfordjournals.bjc.a046738

Wald, A. (1943). Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. *Transactions of the American Mathematical Society*, *54*(3), 426–482. https://doi.org/10.2307/1990256

Waldron, J. S., Malone, S. M., McGue, M., & Iacono, W. G. (2018). A co-twin control study of the relationship between adolescent drinking and adult outcomes. *Journal of Studies on Alcohol and Drugs*, *79*(4), 635–643. https://doi.org/10.15288/jsad.2018.79.635

Waldron, M., Martin, N. G., & Heath, A. C. (2009). Parental alcoholism and offspring behavior problems: Findings in Australian Children of Twins. *Twin Research and Human Genetics*, *12*(5), 433–440. https://doi.org/10.1375/twin.12.5.433

Walker, S. P., Wachs, T. D., Grantham-McGregor, S., Black, M. M., Nelson, C. A., Huffman, S. L., Baker-Henningham, H., Chang, S. M., Hamadani, J. D., Lozoff, B., Gardner, J. M. M., Powell, C. A., Rahman, A., & Richter, L. (2011). Inequality in early childhood: Risk and protective factors for early child development. *The*

Lancet, *378*(9799), 1325–1338. https://doi.org/10.1016/S0140-6736(11)60555-2

Wallach Kildemoes, H., Toft Sørensen, H., & Hallas, J. (2011). The Danish National Prescription Registry. *Scandinavian Journal of Public Health*, *39*(7_suppl), 38–41. https://doi.org/10.1177/1403494810394717

Waller, R., Hyde, L. W., Klump, K. L., & Burt, S. A. (2018). Parenting is an environmental predictor of callous-unemotional traits and aggression: A monozygotic twin differences study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(12), 955–963. https://doi.org/10.1016/j.jaac.2018.07.882

Warr, M. (1993). Age, Peers, and Delinquency. *Criminology*, *31*(1), 17–40. https://doi.org/10.1111/j.1745-9125.1993.tb01120.x

Weaver, J. M., & Schofield, T. J. (2015). Mediation and moderation of divorce effects on children's behavior problems. *Journal of Family Psychology*, *29*(1), 39–48. https://doi.org/10.1037/fam0000043

Weber, L., Kamp-Becker, I., Christiansen, H., & Mingebach, T. (2019). Treatment of child externalizing behavior problems: A comprehensive review and meta–meta-analysis on effects of parent-based interventions on parental characteristics. *European Child & Adolescent Psychiatry*, *28*(8), 1025–1036. https://doi.org/10.1007/s00787-018-1175-3

Wellcome. (2021, August). *Clinical trials policy – Grant funding*. Wellcome. https://wellcome.org/grant-funding/guidance/clinical-trials-policy

Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2000). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Oxford.

Widom, C. S., & Wilson, H. W. (2015). Intergenerational Transmission of Violence. In J. Lindert & I. Levav (Eds.), *Violence and Mental Health: Its Manifold Faces* (pp. 27–45). Springer Netherlands. https://doi.org/10.1007/978-94-017-8999-8_2

Williams, L. D., & Lawrence Aber, J. (2016). Testing for plausibly causal links between parental bereavement and child socio-emotional and academic outcomes: A propensity-score matching model. *Journal of Abnormal Child Psychology*, *44*(4), 705–718. https://doi.org/10.1007/s10802-015-0069-9

Williamson, H. C., Bradbury, T. N., Trail, T. E., & Karney, B. R. (2011). Factor Analysis of the Iowa Family Interaction Rating Scales. *Journal of Family Psychology : JFP : Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *25*(6), 993–999. https://doi.org/10.1037/a0025903

Wilson, H. W., Stover, C. S., & Berkowitz, S. J. (2009). Research Review: The relationship between childhood violence exposure and juvenile antisocial

behavior: a meta-analytic review. *Journal of Child Psychology and Psychiatry*, *50*(7), 769–779. https://doi.org/10.1111/j.1469-7610.2008.01974.x

World Bank. (2024). *Adjusted net national income per capita (current US$)*. https://data.worldbank.org/indicator/NY.ADJ.NNTY.PC.CD

World Health Organization. (2004). *ICD-10: International statistical classification of diseases and related health problems : tenth revision*. World Health Organization. https://apps.who.int/iris/handle/10665/42980

Wu, Q., White, K. R., & Coleman, K. L. (2015). Effects of kinship care on behavioral problems by child age: A propensity score analysis. *Children and Youth Services Review*, *57*, 1–8. https://doi.org/10.1016/j.childyouth.2015.07.020

Yavorska, O. O., & Burgess, S. (2017). MendelianRandomization: An R package for performing Mendelian randomization analyses using summarized data. *International Journal of Epidemiology*, *46*(6), 1734–1739. https://doi.org/10.1093/ije/dyx034

Zachrisson, H. D., Dearing, E., Lekhal, R., & Toppelberg, C. O. (2013). Little evidence that time in child care causes externalizing problems during early childhood in Norway. *Child Development*, *84*(4), 1152–1170. https://doi.org/10.1111/cdev.12040

Zhao, Q., Wang, J., Hemani, G., Bowden, J., & Small, D. S. (2020). Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *The Annals of Statistics*, *48*(3), 1742–1769. https://doi.org/10.1214/19-AOS1866

Zhu, Z., Wang, X., Li, X., Lin, Y., Shen, S., Liu, C.-L., Hobbs, B. D., Hasegawa, K., Liang, L., Boezen, H. M., Camargo, C. A., Cho, M. H., & Christiani, D. C. (2019). Genetic overlap of chronic obstructive pulmonary disease and cardiovascular disease-related traits: A large-scale genome-wide cross-trait analysis. *Respiratory Research*, *20*(1), 64. https://doi.org/10.1186/s12931-019-1036-8

Zimmerman, A., Garman, E., Avendano-Pabon, M., Araya, R., Evans-Lacko, S., McDaid, D., Park, A.-L., Hessel, P., Diaz, Y., Matijasevich, A., Ziebold, C., Bauer, A., Paula, C. S., & Lund, C. (2021). The impact of cash transfers on mental health in children and young people in low-income and middle-income countries: A systematic review and meta-analysis. *BMJ Global Health*, *6*(4), e004661. https://doi.org/10.1136/bmjgh-2020-004661

Zvara, B. J., Burchinal, M., & the Family Life Project Key Contributors. (2019). Maternal history of childhood maltreatment and children's cognitive and social development. *Early Child Development and Care*, *191*(3), 403–414. https://doi.org/10.1080/03004430.2019.1621861

Zvara, B. J., Mills-Koonce, R., Carmody, K. A., & Cox, M. (2017). Maternal childhood sexual trauma and child conduct problems. *Journal of Family Violence*, *32*(2), 231–242. https://doi.org/10.1007/s10896-016-9876-1

Zvara, B. J., Sheppard, K. W., & Cox, M. (2018). Bidirectional effects between parenting sensitivity and child behavior: A cross-lagged analysis across middle childhood and adolescence. *Journal of Family Psychology*, *32*(4), 484–495. https://doi.org/10.1037/fam0000372

Zylbersztejn, A., Lewis, K., Nguyen, V., Matthews, J., Winterburn, I., Karwatowska, L., Barnes, S., Lilliman, M., Saxton, J., Stone, A., Boddy, K., Downs, J., Logan, S., Rahi, J., Black-Hawkins, K., Dearden, L., Ford, T., Harron, K., De Stavola, B., & Gilbert, R. (2023). Evaluation of variation in special educational needs provision and its impact on health and education using administrative records for England: Umbrella protocol for a mixed-methods research programme. *BMJ Open*, *13*(11), e072531. https://doi.org/10.1136/bmjopen-2023-072531

Appendix A. Protocol (Chapter 2)

Chapter 2 is based on an article published in the *BMJ Open:*

Karwatowska, L., Russell, S., Solmi, F., Stavola, B. L. D., Jaffee, S., Pingault, J.-B., & Viding, E. (2020). Risk factors for disruptive behaviours: Protocol for a systematic review and meta-analysis of quasi-experimental evidence. BMJ Open, 10(9), e038258. https://doi.org/10.1136/bmjopen-2020-038258

A PDF of this article is included below.

Table 1 Search strategy used in the database searches

Search terms	
Database	**MeSH terms**
Ovid MEDLINE	**Causal inference methods** – causality/; adoption/; child, adopted/; exp twins/; twin study/; propensity score/; siblings/; interrupted time series analysis/; mendelian randomization analysis/; ecological momentary assessment/; fertilization in vitro/; controlled before-after studies/; fuzzy logic/. **Disruptive behaviours** – conduct disorder/; "attention deficit and disruptive behavior disorders"/; antisocial personality disorder/.
Ovid EMBASE	**Causal inference methods** – causality/; causal attribution/; causal modelling/; quasi experimental study/; adopted child/; adoption/; twins/; twin study/; propensity score/; sibling/; sibling relation/; instrumental variable analysis/; time series analysis/; mendelian randomization analysis/; ecological momentary assessment/; in vitro fertilization/; exogenous variable/; fuzzy logic/; fuzzy system/; maximum likelihood method/. **Disruptive behaviours** – conduct disorder/; oppositional defiant disorder/; disruptive behavior/; antisocial personality disorder/; psychopathy/.
Ovid PsycINFO	**Causal inference methods** – exp causality/; exp causal analysis/; exp quasi experimental methods/; adopted children/; “adoption (child)”/; adoptive parents/; twins/; exp siblings/; exp time series/; exp ecological momentary assessment/; reproductive technology/; counterfactual thinking/; fuzzy logic/; fuzzy set theory/; exp maximum likelihood/. **Disruptive behaviours** – exp conduct disorder/; exp oppositional defiant disorder/; exp disruptive behavior disorders/; exp externalizing symptoms/; antisocial personality disorder/; psychopathy/.
Web of Science	Causal inference methods – Not applicable. Disruptive behaviours – Not applicable.
Free text terms	
Concept 1 - Causal inference methods	Causal inference methods MeSH Terms (causal*) ((quasiexperiment*) or (quasi experiment*)) (adopt*) (fixed effect*) (twin*) (propensity score*) (sibling*)

	(regression discontinuity) (instrumental variable*) (interrupted time series) (mendelian randomi?ation) (matching stud*) (experience sampl*) (ecological momentary assessment*) ((difference* in difference*) or (difference* stud*)) (in vitro fertili?ation) ((polygenic score*) or (polygenic risk score*)) (exogenous varia*) (natural experiment*) (matched control*) (counterfactual*) (potential outcome*) ((balancing adj3 covariate) or (imbalance adj3 covariate) or (balanced adj3 covariate) or (imbalanced adj3 covariate)) (controlled before and after) or (controlled before after)) (inverse probability weight*) ((doubly robust regression*) or (doubly robust estimate*)) ((selection model*) or (selectivity model*)) ((heckit model*) or (heckman sample selection*)) (selection correction*) (two stage residual inclusion*) ((sharp design*) or (fuzzy design*)) (forcing variable*) (full information maximum likelihood) (natural control*) 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31 OR 32 OR 33 OR 34 OR 35
Concept 2 – Disruptive behaviours	Disruptive behaviours MeSH Terms (conduct problem*) (conduct disorder*) (oppositional defian*) (disruptive behavior*) or (disruptive behaviour*) (externali?ing) (antisocial personalit*) or (anti social personalit*) (dissocial personalit*) ((psychopathic) or (psychopathy) or (psychopath) or (psychopaths)) 37 OR 38 OR 39 OR 40 OR 41 OR 42 OR 43 OR 44 OR 45
	36 AND 46 Limit to English Language Human not animal Not review or meta-analysis or case-report Limit to 1980 - Current

APPENDIX A – PROTOCOL

Table 2 Tools and techniques for searching the databases.

Technique and description	Command	Example
All known synonyms and spellings of key words		Doubly robust regression could also be referred to as doubly robust estimate
Replace up to one character in the word – allows alternative spellings to be included.	?	Mendelian randomi?ation would include Mendelian randomisation and Mendelian randomization
Truncation command – used to acknowledge and capture alternative endings to words.	*	Inverse probability weight* would additionally search for inverse probability weights and inverse probability weighting
Boolean logic operators - used to either: a) identify results with at least one of the search terms present; and b) to combine results of different search terms.	"OR" "AND"	Conduct problem* OR conduct disorder* would retrieve articles that have either terms. *Conduct problem** *AND causal** would only retrieve articles that have *both* terms.
Proximity operators - used to identify words within a specified distance of each other.	adj3	*Balancing adj3 covariate* would identify articles whereby "balancing" and "covariate" are within three words of each other.

Table 3 The original Newcastle-Ottawa Scale for Cohort Studies (right) and the adapted version (left) that we used in the current study. Changes are shown in blue.

Original Newcastle-Ottawa Scale	Adapted Newcastle-Ottawa Scale
Selection	
1) Representativeness of the exposed cohort a) truly representative of the average _____ (describe) in the community - b) somewhat representative of the average _____ in the community - c) selected group of users e.g. nurses, volunteers d) no description of the derivation of the cohort	1) Representativeness of the exposed cohort a) truly representative of the average cohort in the community (1) b) somewhat representative of the average cohort in the community (0.5) c) selected group of users e.g. nurses, volunteers (0) d) no description of the derivation of the cohort (0)
2) Selection of the non-exposed cohort a) drawn from the same community as the exposed cohort - b) drawn from a different source c) no description of the derivation of the non-exposed cohort	2) Selection of the non-exposed cohort a) drawn from the same community as the exposed cohort (1) b) drawn from a different source (0) c) no description of the derivation of the non-exposed cohort (0)
3) Ascertainment of exposure a) secure record (e.g. surgical records) - b) structured interview - c) written self-report d) no description	**3) Ascertainment of exposure** **a) validated measure (1)** **b) non-validated measure or no description (0)**
4) Demonstration that outcome of interest was not present at start of study a) yes - b) no	4) Demonstration that outcome of interest was not present **prior to exposure, or control for pre-existing outcome** a) yes (1) b) no (0)
Comparability	
1) Comparability of cohorts on the basis of the design or analysis a) study controls for _____ (select the most important factor) -	**1) Study accounts for the majority of environmental confounders, either by design or statistically accounting for wide range of measured variables** **a) yes (e.g., QE study controlling for other factors) (1)**

b) study controls for any additional factor ¯ (This criteria could be modified to indicate specific control for a second important factor.)	**b) some but not all (e.g., QE study controlling for a few other factors) (0.5)** **c) no (0)** **2) Study fully accounts for genetic confounding** **a) yes (e.g., MZ twin design) (1)** **b) somewhat (e.g. DZ twin design, sibling design, or control for polygenic score or family history of outcome) (0.5)** **c) no (0)**
Outcome	
1) Assessment of outcome a) independent blind assessment ¯ b) record linkage ¯ c) self-report d) no description	1) Assessment of outcome **a) validated measure (1)** **b) non-validated measure or no description (0)** **2) Exposure and outcome reported by different informants** **a) yes (1)** **b) no (0)**
2) Was follow-up long enough for outcomes to occur a) yes (select an adequate follow up period for outcome of interest) ¯ b) no	**3) Exposure and outcome were assessed at the same time** **a) no (assessment was longitudinal – i.e., after exposure) (1)** **b) yes – cross-sectional study / outcome assessed concurrently (0)**
3) Adequacy of follow up of cohorts a) complete follow up - all subjects accounted for ¯ b) subjects lost to follow up unlikely to introduce bias - small number lost - > ___ % (select an adequate follow up, or description provided of those lost) c) follow up rate < ___% (select an adequate, or no description of those lost) d) no statement	4) Adequacy of follow up of cohorts a) complete follow up - all subjects accounted for (1) b) subjects lost to follow up unlikely to introduce bias - small number lost - > 70 % follow up, or method to account for attrition employed) (1) c) follow up rate < 70% and no description of those lost d) no statement (0)

Appendix B. Systematic Review (Chapter 3)

Table 1. Summary of all the studies included in the systematic review of risk factors for disruptive behaviour disorders.

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Anderson & Leventhal (2017)	NICHD	USA	Propensity score matching	Residential mobility	N/A	Externalising problems	1,056
Anthony et al., (2019)	WACS	UK	Adoption study	Parenting	Warmth	Externalising problems	62
Asbury et al., (2003)	TEDS	UK	Discordant twin study	Parenting	Harsh discipline and negative parental feeling	Other DBD	4,706
Asbury et al., (2006)	TEDS	UK	Discordant twin study	Parenting	Discipline, negative parental feelings, parent-child communication	Other DBD	5,162
Asbury et al., (2006)	TEDS	UK	Discordant twin study	Birth weight	N/A	Other DBD	5,162
Averdijk et al., (2018)	Z-PROSO	Switzerland	Propensity score matching	Adverse childhood experiences	Foster care	Externalising problems	1,483
Barnett & Scaramella (2013)	Not reported	Not reported	Discordant sibling study	Parenting	Positive and negative	Externalising problems	274
Beach et al., (2013)	Iowa Adoptee Study	USA	Adoption study	Adverse childhood experiences	Sexual abuse	Antisocial personality disorder	155

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Berger et al., (2005)	NLSCY	USA	Propensity score matching	Parental education and employment	Maternity leave	Externalising problems	769
Besemer et al., (2016)	PYS	USA	Within-individual fixed effects	Parenting	Harsh discipline, involvement, communication	Conduct problems or oppositional defiant disorder	487
Boisvert & Wright (2008)	PSID	USA	Discordant sibling study	Parenting	Warmth, monitoring	Externalising problems	1,759
Boisvert & Wright (2008)	PSID	USA	Discordant sibling study	Peer relationships	Delinquency	Externalising problems	1,759
Boutwell & Beaver (2010)	ECLS-B	USA	Propensity score matching	Prenatal exposure	Smoking	Externalising problems	3,343
Boutwell et al., (2011)	FFCWS	USA	Propensity score matching	Prenatal exposure	Smoking	Externalising problems	1,951
Boyle et al., (2004)	OCHS	Canada	Discordant sibling study	Parenting	Positive and hostile	Externalising problems	2,128
Boyle et al., (2004)	NLSCY	USA	Discordant sibling study	Parenting	Positive and hostile	Externalising problems	7,392
Boyle et al., (2004)	NLSY79	USA	Discordant sibling study	Parenting	Positive and hostile	Externalising problems	2,876
Bradshaw et al., (2020)	GUI	Ireland	Propensity score matching	Parental externalising	Incarceration	Externalising problems	100
Brandlistuen et al., (2013)	MoBa	Norway	Discordant sibling study	Prenatal exposure	Paracetamol	Externalising problems	1,878
Brandlistuen et al., (2015)	MoBa	Norway	Discordant sibling study	Prenatal exposure	Antidepressants	Externalising problems	20,180
Brandlistuen et al., (2017)	MoBa	Norway	Discordant sibling study	Prenatal exposure	Antidepressants	Externalising problems	38,594
Brunborg et al., (2014)	Young in Norway	Norway	Within-individual fixed effects	Gaming	N/A	Conduct problems	1,928

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Bubonya et al., (2019)	LSAC	Australia	Difference in difference	Neighbourhood characteristics	Employment	Externalising problems	4,089
Burt et al., (2005)	MTFS	USA	Adoption study	Parenting	Parent-child relationship	Conduct disorder and oppositional defiant disorder	1,506
Burt et al., (2008)	SIBS	USA	Adoption study	Family dynamics	Separation	Other DBD	204
Burt et al., (2009)	MTFS	USA	Discordant twin study	Peer relationships	Affiliation	Externalising behaviours, conduct disorder and oppositional defiant disorder	908
Burt et al., (2010)	MTFS	USA	Discordant twin study	Family dynamics	Separation	Antisocial personality disorder	578
Cadoret (1995)	Not reported	Not reported	Adoption study	Adverse childhood experiences	ACEs	Conduct disorder	197
Caspi et al., (2004)	E-Risk	UK	Discordant twin study	Parenting	Expressed emotion	Externalising problems	1,244
Cecil et al., (2012)	TEDS	UK	Discordant twin study	Parenting	Harsh discipline and negative parental feeling	Conduct problems	5,184
Chao et al., (2017)	BeTwiSt	China	Mendelian randomisation	Own substance use	Alcohol	Externalising problems	1,608
Copp et al., (2018)	FFCWS	USA	Propensity score matching	Parental externalising	Incarceration	Externalising problems	3,196

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Costello et al., (2003)	GSMS	USA	Natural experiment	Socio-economic position	N/A	Conduct disorder or oppositional defiant disorder	1,420
Costello et al., (2010)	GSMS	USA	Natural experiment	Socio-economic position	N/A	Conduct disorder, oppositional defiant disorder, or antisocial personality disorder	1,420
Crosby et al., (2010)	Six random-assignment studies	Various	Instrumental variable analyses	Own education	Childcare	Externalising problems	3,290
D'Onofrio et al., (2008)	NLSY79	USA	Discordant sibling study	Prenatal exposure	Smoking	Conduct problems	6,283
D'Onofrio et al., (2009)	NLSY79	USA	Discordant sibling study	Maternal age at birth	N/A	Conduct problems	15,763
D'Onofrio et al., (2009)	NLSY79	USA	Discordant sibling study	Socio-economic position	N/A	Conduct problems	4,912
D'Onofrio et al., (2012)	NLSY79	USA	Discordant sibling study	Prenatal exposure	Smoking	Externalising problems	6,066
D'Onofrio, Slutske et al., (2007)	ATR	Australia	Children of twins study	Parental externalising	Conduct disorder	Conduct disorder	2,554
D'Onofrio, Van Hulle et al., (2007)	NLSY79	USA	Discordant sibling study	Prenatal exposure	Alcohol	Externalising problems	3,447
Deater-Deckard & Petrill, (2004)	N2CAP	USA	Adoption study	Parenting	Mutuality	Externalising problems	396

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Dee & Sievertsen (2018)	DNBC	Denmark	Fuzzy regression discontinuity design	Own education	School starting age	Conduct problems	7,642
Dinwiddie et al., (2000)	ATR	Australia	Discordant twin study	Adverse childhood experiences	Sexual abuse	Conduct problems	2,682
Doi et al., (2018)	A-CHILD	Japan	Propensity score matching	Adverse childhood experiences	Neglect	Externalising behaviour	4,195
Dunifon et al., (2003)	WES	USA	Within-individual fixed effects	Parental education and employment	Status	Externalising problems	573
Duyme (1990)	Not reported	France	Adoption study	Parental education and employment	Type	Other DBD	77
D'Onofrio et al., (2005)	OZALC	Australia	Children of twins study	Family dynamics	Separation	Externalising problems	2,554
Edwards & Yu (2018)	YPCA	Australia	Propensity score matching	Own education	Childcare	Externalising problems	317
Ekblad et al., (2017)	Finnish Medical Birth Register	Finland	Discordant sibling study	Prenatal exposure	Smoking	Externalising problems	300,336
Ekblad et al., (2020)	MO-MATCH	USA	Discordant sibling study	Prenatal exposure	Smoking	Externalising problems	346
Ellingson et al., (2014)	NLSY79	USA	Discordant sibling study	Prenatal exposure	Smoking	Externalising problems	10,251
Ellis et al., (2012)	TESS	Norway	Propensity score matching	Prenatal exposure	Smoking	Oppositional defiant disorders	995

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Emery, (2011)	PHDCN	USA	Within-individual fixed effects	Adverse childhood experiences	Intimate partner violence	Externalising behaviour	1,816
Estabrook et al., (2016)	MIDS	USA	Discordant sibling study	Prenatal exposure	Smoking	Externalising problems	299
Fitzsimons & Villadsen (2019)	MCS	UK	Within-individual fixed effects	Family dynamics	Separation	Conduct problems	6,245
Gaysina et al., (2013)	CHDS	New Zealand	Adoption study	Prenatal exposure	Smoking	Conduct problems	36
Gaysina et al., (2013)	EGDS	USA	Adoption study	Prenatal exposure	Smoking	Conduct problems	311
Gaysina et al., (2013)	C-IVF	UK	IVF study	Prenatal exposure	Smoking	Conduct problems	206
Gershoff et al., (2018)	ECLS-K	USA	Propensity score matching	Adverse childhood experiences	Physical abuse	Externalising problems	12,112
Gilman et al., (2008)	CPP	USA	Discordant sibling study	Prenatal exposure	Smoking	Conduct problems	52,919
Girard et al., (2018)	GUI	Ireland	Propensity score matching	Breastfeeding	N/A	Externalising problems	6,013
Girard & Farkas (2019)	ELPI	Chile	Propensity score matching	Breastfeeding	N/A	Externalising problems	3,037
Gjerde et al., (2017)	MoBa	Norway	Discordant sibling study	Parental internalising	Depression	Externalising problems	17,830
Gjerde et al., (2020)	MoBa	Norway	Discordant sibling study	Parental internalising	Anxiety	Externalising problems	17,724
Glover et al., (2010)	N2CAP	USA	Adoption study	Parenting	Positive and negative	Externalising problems	85
Goldberg & Carlson (2014)	FFCWS	USA	Within-individual fixed effects	Family dynamics	Parental relationship quality	Externalising problems	773

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Gomajee et al., (2018)	EDEN	France	Propensity score matching	Own education	Childcare	Other DBD	1,428
Goodnight et al., (2016)	NLSY79	USA	Discordant sibling study	Temperament	N/A	Conduct problems and delinquency	9,237
Grabow et al., (2017)	EGDS	USA	Adoption study	Parental internalising	Depression	Externalising problems	541
Grabow et al., (2017)	EGDS	USA	Adoption study	Parental adverse childhood experience	Trauma	Externalising problems	541
Haber et al., (2005)	VET	USA	Children of twins study	Parental substance use	Alcohol	Conduct disorder	1,270
Haber et al., (2010)	VET	USA	Children of twins study	Parental substance use	Alcohol and drugs	Conduct disorder	1,917
Hails et al., (2019)	EGDS	USA	Adoption study	Parental internalising	Depression	Externalising problems	503
Hannigan et al., (2018)	MoBa	Norway	Children of twins study	Parental internalising	Post-natal depression	Externalising problems	35,299
Harden et al., (2009)	NLSY79	USA	Within-individual fixed effects	Neighbourhood characteristics	Density	Conduct problems	9,440
Harden, Lynch et al., (2007)	OZALC	Australia	Children of twins study	Maternal age at birth	N/A	Other DBD	1,364
Harden , Turkheimer et al., (2007)	OZALC	Australia	Children of twins study	Family dynamics	Marital conflict	Conduct problems	1,131
Harding (2015)	HSIS	UK	Propensity score matching	Parental education and employment	Parental education	Externalising problems	1,588
Harold et al., (2011)	C-IVF	UK	IVF study	Parental externalising	Antisocial behaviour	Other DBD	283

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Harold et al., (2012)	C-IVF	UK	IVF study	Parenting	Hostile	Conduct problems	377
Harold et al., (2013)	EGDS	USA	Adoption study	Parenting	Hostile	Externalising problems	218
Harris et al., (2018)	MoBa	Norway	Inverse probability of treatment weighting	Prenatal exposure	Triptans	Externalising problems	37,656
Haskins (2015)	FFCWS	USA	Propensity score matching	Parental externalising	Incarceration	Externalising problems	2,162
Herbst & Tekin (2016)	ECLS-K	USA	Instrumental variable analyses	Own education	Childcare	Externalising problems	3,848
Hipwell et al., (2016)	PGS	USA	Propensity score matching	Adolescent childbearing	N/A	Conduct problems	441
Hou et al., (2013)	BeTwiSt	China	Discordant twin study	Parenting	Warmth and hostile	Externalising problems	1,040
Humphrey & Root (2017)	ECLS-K	USA	Propensity score matching	Neighbourhood characteristics	Disadvantage	Externalising problems	14,960
Ichikawa et al., (2017)	J-SHINE	Japan	Within-individual fixed effects	Neighbourhood characteristics	Social cohesion and informal social control	Externalising problems	918
Jackson (2018)	ECLS-B	USA	Propensity score matching	TV viewing	N/A	Externalising problems	5,000
Jaffee et al., (2011)	CNLSY	USA	Discordant sibling study	Own education	Childcare	Conduct problems	9,185
Kendler et al., (2020)	Swedish population-based registers	Sweden	Discordant sibling study	Parental substance use	Alcohol and drugs	Conduct disorder	146,216

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Kendler et al., (2020)	Swedish population-based registers	Sweden	Discordant sibling study	Parental internalising	Depression	Conduct disorder	146,216
Kerr et al., (2013)	EGDS	USA	Adoption study	Parental internalising	Depression	Other DBD	346
Kerr et al., (2013)	EGDS	USA	Adoption study	Parental externalising	Antisocial behaviour	Other DBD	347
Keyes et al., (2008)	SIBS	USA	Adoption study	Parental substance use	Smoking	Other DBD	785
King et al., (2009)	SIBS	USA	Adoption study	Parental substance use	Alcohol	Other DBD	525
King (2018)	FFCWS	USA	Within-individual fixed effects	Socio-economic position	Poverty	Externalising behaviour	2,044
King (2018)	FFCWS	USA	Within-individual fixed effects	Parental internalising	Depression	Externalising behaviour	2,044
King (2018)	FFCWS	USA	Within-individual fixed effects	Parental internalising	Stress	Externalising behaviour	2,044
Klahr, McGue et al., (2011)	SIBS	USA	Adoption study	Parenting	Parent-child relationship	Conduct problems and antisocial personality disorder	672
Klahr, Rueter et al., (2011)	SIBS	USA	Adoption study	Parenting	Parent-child conflict and coercive	Other DBD	390
Knudsen et al., (2015)	MoBa	Norway	Within-individual fixed effects	Parental substance use	Alcohol	Externalising problems	51,115
Latham et al., (2017)	TFaB	UK	Discordant sibling study	Parenting	Coercive	Externalising problems	212

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Lee & McLanahan (2015)	FFCWS	USA	Inverse probability of treatment weighting	Family dynamics	Separation	Externalising problems	2,952
Lee et al., (2018)	NICHD SECCYD	USA	Within-individual fixed effects	Own education	Unsupervised time with peers	Externalising problems	747
Lee et al., (2018)	NICHD SECCYD	USA	Within-individual fixed effects	Own education	Paid employment	Externalising problems	747
Lee et al., (2018)	NICHD SECCYD	USA	Within-individual fixed effects	Own education	Sports	Externalising problems	747
Lee et al., (2018)	NICHD SECCYD	USA	Within-individual fixed effects	Own education	Other organised activities after-school	Externalising problems	747
Leve et al., (2009)	EGDS	USA	Adoption study	Parenting	Structured	Externalising problems	290
Lipscomb et al., (2014)	EGDS	USA	Adoption study	Parenting	Overreactive	Other DBD	233
Lipscomb et al., (2014)	EGDS	USA	Adoption study	Own education	Childcare	Other DBD	233
Lombardi & Coley (2014)	ECLS–B	USA	Propensity score matching	Parental education and employment	Intensity	Conduct problems	10,100
Long et al., (2015)	VATSPSUD	USA	Discordant twin study	Parenting	Care and overprotection	Conduct disorder	2,606
Lund et al., (2019)	MoBa	Norway	Instrumental variable analyses	Parental substance use	Alcohol	Externalising problems	25,744
Lysenko et al., (2013)	TEDS	UK	Discordant sibling study	Parenting	Harsh discipline	Conduct problems	27,660
Ma et al., (2018)	FFCWS	USA	Within-individual fixed effects	Neighbourhood characteristics	Disorganisation	Other DBD	2,472
Ma et al., (2018)	FFCWS	USA	Within-individual fixed effects	Adverse childhood experiences	Physical abuse	Other DBD	2,472

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Ma et al., (2020)	FFCWS	USA	Within-individual fixed effects	Neighbourhood characteristics	Violence	Externalising problems	2,472
Ma et al., (2020)	FFCWS	USA	Within-individual fixed effects	Adverse childhood experiences	Physical abuse	Externalising problems	2,472
Mankuta et al., (2010)	LIST	Israel	Discordant twin study	Birth weight	N/A	Conduct problems	224
Marceau et al., (2013)	EGDS	USA	Adoption study	Parenting	Overreactive	Externalising problems	561
Marceau et al., (2013)	EGDS	USA	Adoption study	Parental internalising	Depression and anxiety	Externalising problems	561
Mark & Pike (2017)	SBS	UK	Discordant sibling study	Parenting	Parent-child relationship	Conduct problems	156
McAdams et al., (2015)	TOSS	Sweden	Children of twins study	Parental internalising	Depression	Externalising problems	1,752
McAdams et al., (2015)	EGDS	USA	Adoption study	Parental internalising	Depression	Externalising problems	361
McCrory & Layte (2012)	GUI	Ireland	Propensity score matching	Prenatal exposure	Smoking	Externalising problems	7,505
Meunier et al., (2012)	HBHC	Canada	Discordant sibling study	Parenting	Positive and negative	Externalising problems	809
Misheva et al., (2017)	ATR	Australia	Discordant twin study	Adverse childhood experiences	Sexual abuse and physical maltreatment	Conduct disorder	11,060
Møllegaard (2020)	DMTS	Denmark	Discordant twin study	Birth weight	N/A	Externalising problems	4,228
Monnet (2019)	NSCH	USA	Instrumental variable analyses	Own education	Childcare	Conduct problems	42,462
Morcillo et al., (2011)	BYS	USA/Puerto Rico	Propensity score matching	Parenting	Parent-child relationship	Other DBD	2,491

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Mostafa et al., (2018)	MCS	UK	Within-individual fixed effects	Family dynamics	Separation	Externalising problems	14,833
Murray et al., (2016)	ALSPAC	UK	Mendelian randomisation	Prenatal exposure	Alcohol	Other DBD	3,544
Narusyte et al., (2011)	TOSS & TCHAD	Sweden	Children of twins study	Parenting	Criticism	Externalising problems	1,818
Neiderhiser et al., (2016)	EGDS	USA	Adoption study	Obstetric complications	N/A	Externalising problems	561
Neiderhiser et al., (2016)	EGDS	USA	Adoption study	Prenatal exposure	Substances and toxins	Externalising problems	561
Nelson et al., (2002)	ATR	Australia	Discordant twin study	Adverse childhood experiences	Sexual abuse	Conduct disorder	1,991
Neugebauer et al., (1999)	Not reported	The Netherlands	Natural experiment	Prenatal exposure	Nutrient deficiency	Antisocial personality disorder	100,543
Nulman et al., (2015)	Toronto Sibling Study / Motherisk	Canada	Discordant sibling study	Prenatal exposure	Antidepressants	Externalising problems	90
	Z-PROSO	Switzerland	Propensity score matching	Own education	Teacher-student relationship	Other DBD	1,067
Oliver et al., (2008)	TEDS	UK	Discordant twin study	Own education	Classroom environment	Externalising problems	570
Oliver (2015)	TEDS	UK	Discordant twin study	Parenting	Negative	Conduct problems	6,308
Orri et al., (2019)	QLSCD	Canada	Propensity score matching	Own education	Childcare	Conduct problems	1,588
Palmer et al., (2016)	MOAFTS	USA	Propensity score matching	Prenatal exposure	Smoking	Externalising problems	3,232

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Paradis et al., (2017)	CPP	USA	Discordant sibling study	Prenatal exposure	Smoking	Other DBD	1,684
Pemberton et al., (2010)	EGDS	USA	Adoption study	Parental internalising	Depression	Externalising problems	351
Pike et al., (1996)	NEAD	USA	Discordant twin study	Parenting	Negative	Other DBD	93
Powers et al., (2016)	Fast Track Project	USA	Propensity score matching	Own education	Special education setting	Conduct disorder	891
Ramanathan et al., (2017)	NLSY79	USA	Discordant sibling study	Socio-economic position	N/A	Externalising problems	8,276
Reuben et al., (2016)	EGDS	USA	Adoption study	Parenting	Warmth and overreactive	Externalising problems	225
Rice et al., (2010)	C-IVF	UK	IVF study	Parental internalising	Stress	Other DBD	474
Richmond & Stocker (2003)	Not reported	USA	Discordant sibling study	Family dynamics	Marital conflict	Externalising problems	122
Richmond & Stocker (2006)	Not reported	Not reported	Discordant sibling study	Parenting	Hostile	Externalising problems	186
Richmond & Stocker (2009)	Not reported	Not reported	Discordant sibling study	Parenting	Hostile	Other DBD	228
Richmond & Stocker (2009)	Not reported	Not reported	Discordant sibling study	Family dynamics	Separation	Other DBD	228
Riggins-Caspers et al., (2003)	Not reported	USA	Adoption study	Parenting	Harsh discipline	Conduct disorder	150
Rivenbark et al., (2020)	E-Risk	UK	Discordant twin study	Socio-economic position	N/A	Conduct problems	2,232

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Rolon-Arroyo et al., (2018)	Not reported	USA	Within-individual fixed effects	Parenting	Positive and negative and warmth	Conduct disorder and oppositional defiant disorder	199
Roos et al., (2016)	EGDS	USA	Adoption study	Parenting	Uninvolved	Externalising problems	293
Roos et al., (2016)	EGDS	USA	Adoption study	Parental internalising	Internalising	Externalising problems	293
Rudolph et al., (2019)	NCS-A	USA	Propensity score matching	Neighbourhood characteristics	Noise	Other DBD	2,645
Samek et al., (2014)	SIBS	USA	Adoption study	Parenting	Parent-child relationship	Antisocial personality disorder	533
Samek et al., (2014)	SIBS	USA	Adoption study	Parental substance use	Alcohol and smoking	Antisocial personality disorder	533
Sampson & Winter (2018)	PHDCN	USA	Propensity score matching	Prenatal exposure	Lead exposure	Externalising problems	1,255
Sampson & Winter (2018)	PHDCN	USA	Instrumental variable analyses	Prenatal exposure	Lead exposure	Externalising problems	1,255
Schermerhorn et al., (2011)	TOSS	Sweden	Children of twins study	Family dynamics	Family functioning	Externalising problems	1,818
Scherrer et al., (1996)	VET	USA	Discordant twin study	Parental substance use	Cannabis	Conduct disorder and antisocial personality disorder	3,394
Shelton et al., (2008)	CaStANET	UK	Discordant twin study	Parenting	Hostile and warmth	Conduct problems	462
Shelton et al., (2011)	Not reported	USA/UK	IVF study	Breastfeeding	N/A	Conduct problems	870

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Singh et al., (2011)	ATR	Australia	Children of twins study	Parental internalising	Depression	Conduct disorder	2,554
Singham et al., (2017)	TEDS	UK	Discordant twin study	Peer relationships	Bullying	Conduct problems	11,108
Skopp et al., (2005)	Not reported	USA	Discordant sibling study	Family dynamics	Marital conflict	Externalising problems	244
Sonuga-Barke et al., (2017)	ERA	UK	Natural experiment	Adverse childhood experiences	Deprivation	Other DBD	217
Staff et al., (2019)	MCS	UK	Within-individual fixed effects	Own substance use	Alcohol	Externalising problems	10,529
Sundbakk et al., (2019)	MoBa	Norway	Inverse probability of treatment weighting	Prenatal exposure	Antidepressants	Externalising problems	36,401
Taraban et al., (2019)	EGDS	USA	Adoption study	Parental internalising	Depression	Externalising problems	1,038
Tore et al., (2018)	TAMBAHS	UK	Discordant twin study	Birth weight	N/A	Externalising problems	508
Trønnes et al., (2020)	MoBa	Norway	Inverse probability of treatment weighting	Prenatal exposure	Paracetamol	Externalising problems	32,934
Tully et al., (2008)	SIBS	USA	Adoption study	Parental internalising	Depression	Conduct disorder	568
Turney (2012)	FFCWS	USA	Propensity score matching	Parental internalising	Depression	Behavioural problems	2,655
Turney (2017)	FFCWS	USA	Propensity score matching	Parental externalising	Incarceration	Externalising problems	3,065

Reference	Cohort	Country	Method	Risk factor category	Risk factor subcategory	Outcome	Sample size
Viding et al., (2009)	TEDS	UK	Discordant twin study	Parenting	Discipline	Conduct problems and callous–unemotional traits	4,508
Waldron et al., (2009)	OZALC	Australia	Children of twins study	Parental substance use	Alcohol	Externalising problems	2,492
Waldron et al., (2018)	MTFS	USA	Discordant twin study	Own substance use	Alcohol	Other DBD	2,764
Waller et al., (2018)	TBED-C	USA	Discordant twin study	Parenting	Warmth and harsh discipline	Aggression and callous-unemotional traits	454
Weaver & Schofield, (2015)	NICHD	USA	Propensity score matching	Family dynamics	Separation	Externalising problems	1,364
Williams & Lawrence Aber, (2016)	ECLS-K	USA	Propensity score matching	Adverse childhood experiences	Parental bereavement	Externalising problems	250
Wu et al., (2015)	NSCAW II	USA	Propensity score matching	Adverse childhood experiences	Foster care	Externalising problems	1,054
Zachrisson et al., (2013)	MoBa	Norway	Discordant sibling study	Own education	Childcare	Externalising problems	75,271
Zvara et al., (2017)	FLP	USA	Propensity score matching	Parental adverse childhood experience	Sexual abuse	Conduct problems	204
Zvara et al., (2019)	FLP	USA	Propensity score matching	Parental adverse childhood experience	Sexual abuse	Conduct problems	204

Appendix C. Meta-analysis (Chapter 4)

Table 1. PRISMA reporting checklist.

Note. the page numbers relate to the manuscript not the thesis Chapter.

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	4
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	6
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	11
Search strategy	7	Present the full search strategies for all databases, registers, and websites, including any filters and limits used.	Table S4
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	11
Data	9	Specify the methods used to collect data from reports,	11

Section and Topic	Item #	Checklist item	Location where item is reported
collection process		including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	11
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	10
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	12
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	12
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	NA
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Table S6
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	NA

Section and Topic	Item #	Checklist item	Location where item is reported
	13d	Describe any methods used to synthesise results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	12
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	14
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesised results.	13
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	12
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	NA
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Figure S1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	NA
Study characteristics	17	Cite each included study and present its characteristics.	Table 1
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Table 1
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Figures 1 and 2
Results of	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing	Table 3

Section and Topic	Item #	Checklist item	Location where item is reported
syntheses		studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	24 -31
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	26 - 31
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesised results.	26 - 31
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Figures S3 and S4
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	33 - 41
	23b	Discuss any limitations of the evidence included in the review.	40 - 41
	23c	Discuss any limitations of the review processes used.	40 - 41
	23d	Discuss implications of the results for practice, policy, and future research.	39
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	9
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	9
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	9
Support	25	Describe sources of financial or non-financial support	42

Section and Topic	Item #	Checklist item	Location where item is reported
		for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	42
Availability of data, code, and other materials	27	Report which of the following are publicly available and where they can be found template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	9

Table 2. MOOSE Reporting Guidelines

Note. the page numbers relate to the manuscript not the thesis Chapter.

Recommendation	Pg. no.
Reporting background should include	
Problem definition	6
Hypothesis statement	6
Description of study outcome(s)	11
Type of exposure or intervention used	10
Type of study designs used	10
Study population	9-10
Reporting of search strategy should include	
Qualifications of searchers (e.g. librarians and investigators)	11
Search strategy, including time period included in the synthesis and keywords	11
Effort to include all available studies, including contact with authors	11
Databases and registries searched	11
Search software used, name and version, including special features	11
Use of hand searching (e.g. reference lists of obtained articles)	NA
List of citations located and those excluded including justification	NA
Method of addressing articles published in languages other than English	NA (only English language included)
Method of handling abstracts and unpublished studies	NA (only published studies included)
Description of any contact with authors	43
Reporting methods should include	
Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested	10; Table S3
Rationale for the selection and coding of data	9 – 10
Documentation of how data were classified and coded (e.g., multiple raters, blinding, and interrater reliability)	11
Assessment of confounding	13
Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results	12
Assessment of heterogeneity	13-14

Description of statistical methods (e.g., complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated	11
Provision of appropriate tables and graphics	Figures 1-3
Reporting of results should include	
Graphic summarizing individual study estimates and overall estimate	Figures 2 and 3
Table giving descriptive information for each study included	Table 1
Results of sensitivity testing (e.g., subgroup analysis)	Table 4
Indication of statistical uncertainty of findings	25 – 32
Reporting of discussion should include	
Quantitative assessment of bias (e.g., publication bias)	27 - 29
Justification for exclusion (e.g., exclusion of non–English-language citations)	NA
Assessment of quality of included studies	Figures 2 and 3
Reporting of conclusions should include	
Consideration of alternative explanations for observed results	38 - 39
Generalization of the conclusions (i.e., appropriate for the data presented and within the domain of the literature review)	42
Guidelines for future research	40
Disclosure of funding source	42

Table 3. Formulae used to convert original study data into Pearson's r and SE.

Raw effect size type	Formulae
95% confidence intervals	$SE = \frac{uCI - lCI}{3.92}$
Unstandardised beta	$SE = \frac{1 - beta^2}{\sqrt{n-2}}$
Unstandardised beta	$r = \frac{beta * SD^{EXP}}{SD^{OUT}}$

Abbreviations: SE = standard error; uCI = upper 95% confidence interval; lCI = lower 95% confidence interval; SD^{EXP} = standard deviation of the exposure; SD^{OUT} = standard deviation of the outcome.

Table 4. Summary of all the effect sizes included in the meta-analyses of positive parenting practices.

Reference	Country	Method	Sex	Exposure	Age^{EXP}	Outcome	Age^{OUT}	Informant	ES	SE
Anthony (2019)	UK	Adoption study	Mixed sex	Parental Warmth	2.33	Externalising Behaviour	3.66	caregiver/ caregiver	-0.220	0.136
Asbury (2006) [11]	UK	Discordant twin study	Mixed sex (whole sample)	Parent-Child Communication	4.00	Conduct Problems	7.00	caregiver/ teacher	0.080	0.041
Asbury (2006) [12]	UK	Discordant twin study	Mixed sex (whole sample)	Parent-Child Communication	4.00	Conduct Problems	7.00	caregiver/ teacher	-0.010	0.022
Barnett (2013) [1]	USA	Discordant sibling study	Mixed sex (whole sample)	Parental Warmth	2.99	Externalising Behaviour	3.99	observer/ caregiver	-0.018	0.074
Boisvert (2008) [1]	USA	Discordant sibling study	Mixed (whole sample)	Parental Warmth	12.15	Externalising Behaviour	17.15	caregiver/ caregiver	-0.046	0.051
Boisvert (2008) [2]	USA	Discordant sibling study	Mixed (whole sample)	Parental Monitoring	12.15	Externalising Behaviour	17.15	caregiver/ caregiver	-0.116	0.048
Boyle (2004) [4]	Canada	Discordant sibling study	Mixed sex	Parental Involvement	7.47	Externalising Behaviour	7.47	caregiver/ caregiver	0.020	0.009
Boyle (2004) [6]	Canada	Discordant sibling study	Mixed sex	Parental Involvement	7.47	Externalising Behaviour	7.47	teacher/ teacher	0.040	0.016
Boyle (2004) [8]	USA	Discordant sibling study	Mixed sex	Parental Warmth	9.03	Externalising Behaviour	9.03	observer/ caregiver	-0.010	0.022
Caspi (2004) [3]	UK	Discordant twin study	Mixed sex	Expressed Emotion	5.00	Externalising Behaviour	6.50	caregiver/ caregiver	-0.110	0.043
Caspi (2004) [4]	UK	Discordant twin study	Mixed sex	Parental Warmth	5.00	Externalising Behaviour	6.50	caregiver/ caregiver	-0.110	0.043
Caspi (2004) [7]	UK	Discordant twin study	Mixed sex	Expressed Emotion	5.00	Externalising Behaviour	6.50	caregiver/ teacher	-0.120	0.047

Reference	Country	Method	Sex	Exposure	Age^{EXP}	Outcome	Age^{OUT}	Informant	ES	SE
Caspi (2004) [8]	UK	Discordant twin study	Mixed sex	Parental Warmth	5.00	Externalising Behaviour	6.50	caregiver/ teacher	-0.100	0.051
Deater-Deckard (2004)	USA	Adoption study	Mixed sex	Parent-Child Relationship	8.16	Conduct Problems	8.16	observer/ mixed	-0.390	0.118
Glover (2010) [1]	USA	Adoption study	Mixed sex	Expressed Emotion	5.59	Conduct Problems	5.59	caregiver (M)/ caregiver (F)	-0.290	0.112
Glover (2010) [3]	USA	Adoption study	Mixed sex	Expressed Emotion	5.59	Conduct Problems	5.59	caregiver (F)/ caregiver (M)	-0.380	0.115
Hou (2013) [1]	China	Discordant twin study	Mixed sex	Parental Warmth	13.86	Conduct Problems	15.36	mixed/ mixed	0.050	0.054
Hou (2013) [3]	China	Discordant twin study	Mixed sex	Parental Warmth	13.86	Conduct Problems	15.36	mixed/ mixed	-0.080	0.054
Mark (2017) [1]	UK	Discordant sibling study	Mixed sex	Parent-Child Relationship	10.94	Externalising Behaviour	10.94	self/ caregiver	-0.049	0.117
Meunier, Bisceglia (2012) [1]	Canada	Discordant sibling study	Mixed sex	Parent-Child Relationship	4.49	Oppositional Defiant Disorder	4.49	caregiver/ caregiver	-0.310	0.044
Meunier, Bisceglia (2012) [2]	Canada	Discordant sibling study	Mixed sex	Parent-Child Relationship	4.49	Oppositional Defiant Disorder	4.49	caregiver/ caregiver	-0.151	0.059
Morcillo (2011; M) [1]	USA/Puerto Rico	Propensity score matching	Males	Family Bonding	8.00	Conduct Problems	9.00	caregiver/ mixed	-0.036	0.012
Morcillo (2011; M) [2]	USA/Puerto Rico	Propensity score matching	Males	Family Bonding	11.00	Conduct Problems	12.00	caregiver/ mixed	-0.021	0.032

Reference	Country	Method	Sex	Exposure	Age^{EXP}	Outcome	Age^{OUT}	Informant	ES	SE
Morcillo (2011; F) [1]	USA/Puerto Rico	Propensity score matching	Females	Family Bonding	8.00	Conduct Problems	9.00	caregiver/ mixed	-0.059	0.016
Morcillo (2011; F) [2]	USA/Puerto Rico	Propensity score matching	Females	Family Bonding	11.00	Conduct Problems	12.00	caregiver/ mixed	-0.107	0.021
Reuben (2016) [1]	USA	Adoption study	Mixed sex	Parental Warmth	2.25	Externalising Behaviour	6.50	caregiver/ teacher	-0.182	0.066
Reuben (2016) [3]	USA	Adoption study	Mixed sex	Parental Warmth	2.25	Externalising Behaviour	6.50	caregiver/ teacher	0.086	0.070
Waller, Hyde (2018) [1]	USA	Discordant twin study	Mixed sex	Parental Warmth	7.80	Other DBD	7.80	mixed/ caregiver	0.458	0.073
Paine (2021) [1]	UK	Adoption study	Mixed sex	Parental Warmth	4.08	Externalising Behaviour	5.33	caregiver/ caregiver	-0.420	0.085
Shewark (2021) [2]	USA	Adoption study	Mixed	High Parental Warmth	6.00	Externalising Behaviour	7.00	caregiver/ caregiver	-0.130	0.042
Shewark (2021) [4]	USA	Adoption study	Mixed	High Parental Warmth	6.00	Externalising Behaviour	7.00	caregiver/ caregiver	-0.120	0.080

Abbreviations: AgeEXP = offspring age at exposure assessment; AgeOUT = offspring age at outcome assessment; ES = effect size; SE = standard error

Table 5. Summary of all the effect sizes included in the meta-analyses of negative parenting practices.

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{OUT}	Informant	ES	SE
Asbury (2003) [1]	UK	Discordant twin study	Mixed sex	Harsh Discipline	4.00	Conduct Problems	4.00	caregiver/ caregiver	0.150	0.058
Asbury (2003) [2]	UK	Discordant twin study	Mixed sex	Parental Feeling	4.00	Conduct Problems	4.00	caregiver/ caregiver	0.220	0.085
Asbury (2006) [10]	UK	Discordant twin study	Mixed sex (whole sample)	Parental Feeling	4.00	Conduct Problems	7.00	caregiver/ teacher	0.020	0.022
Asbury (2006) [9]	UK	Discordant twin study	Mixed sex (whole sample)	Harsh Discipline	4.00	Conduct Problems	7.00	caregiver/ teacher	0.070	0.036
Barnett (2013) [2]	USA	Discordant sibling study	Mixed sex (whole sample)	Coercive Parenting	2.99	Externalising Behaviour	3.99	observer/ caregiver	0.091	0.118
Besemer (2016) [1]	USA	Within-person fixed effects	Males	Harsh Discipline	8.65	Oppositional Defiant Disorder	10.90	caregiver/mixed	0.023	0.015
Besemer (2016) [2]	USA	Within-person fixed effects	Males	Parental Involvement	8.65	Oppositional Defiant Disorder	10.90	caregiver/mixed	-0.057	0.032
Besemer (2016) [3]	USA	Within-person fixed effects	Males	Parent-Child Communication	8.65	Oppositional Defiant Disorder	10.90	caregiver/mixed	0.045	0.043
Boyle (2004) [1]	Canada	Discordant sibling study	Mixed sex	Harsh Discipline	10.10	Externalising Behaviour	10.10	caregiver/ caregiver	0.210	0.014
Boyle (2004) [2]	Canada	Discordant sibling study	Mixed sex	Harsh Discipline	10.10	Externalising Behaviour	10.10	caregiver/ teacher	0.110	0.014
Boyle (2004) [3]	Canada	Discordant sibling study	Mixed sex	Parental Hostility	7.47	Externalising Behaviour	7.47	caregiver/ caregiver	0.330	0.008
Boyle (2004) [5]	Canada	Discordant sibling study	Mixed sex	Parental Hostility	7.47	Externalising Behaviour	7.47	teacher/ teacher	0.170	0.015

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Boyle (2004) [7]	USA	Discordant sibling study	Mixed sex	Physical Discipline	9.03	Externalising Behaviour	9.03	caregiver/ caregiver	0.130	0.039
Caspi (2004) [1]	UK	Discordant twin study	Mixed sex	Parental Criticism	5.00	Externalising Behaviour	6.50	caregiver/ caregiver	0.160	0.062
Caspi (2004) [2]	UK	Discordant twin study	Mixed sex	Parental Criticism	5.00	Externalising Behaviour	6.50	caregiver/ caregiver	0.160	0.062
Caspi (2004) [5]	UK	Discordant twin study	Mixed sex	Parental Criticism	5.00	Externalising Behaviour	6.50	caregiver/ teacher	0.150	0.058
Caspi (2004) [6]	UK	Discordant twin study	Mixed sex	Parental Criticism	5.00	Externalising Behaviour	6.50	caregiver/ teacher	0.100	0.051
Cecil (2012) [1]	UK	Discordant twin study	Mixed sex	Harsh Discipline	5.75	Conduct Problems	12.00	caregiver/ teacher	0.140	0.056
Cecil (2012) [2]	UK	Discordant twin study	Mixed sex	Parental Feeling	5.75	Conduct Problems	12.00	caregiver/ teacher	0.040	0.023
Glover (2010) [2]	USA	Adoption study	Mixed sex	Parental Feeling	5.59	Conduct Problems	5.59	caregiver (M)/ caregiver (F)	0.320	0.124
Glover (2010) [4]	USA	Adoption study	Mixed sex	Parental Feeling	5.59	Conduct Problems	5.59	caregiver (F)/caregiver (M)	0.470	0.142
Harold (2012) [1]	UK/USA	IVF study	Mixed sex	Parental Hostility	6.72	Conduct Problems	6.72	caregiver/ caregiver	0.240	0.093
Harold (2012) [2]	UK/USA	IVF study	Mixed sex	Parental Hostility	6.72	Conduct Problems	6.72	caregiver/ caregiver	0.350	0.136
Harold (2013) [1]	USA	Adoption study	Mixed sex	Parental Hostility	5.98	Conduct Problems	5.98	caregiver/ caregiver	0.310	0.094

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Harold (2013) [2]	USA	Adoption study	Mixed sex	Parental Hostility	5.98	Conduct Problems	5.98	caregiver/ caregiver	0.210	0.081
Harold (2013) [3]	USA	Adoption study	Mixed sex	Parental Hostility	5.98	Conduct Problems	5.98	caregiver/ caregiver	0.440	0.133
Harold (2013) [4]	USA	Adoption study	Mixed sex	Parental Hostility	5.98	Conduct Problems	5.98	caregiver (F)/caregiver (M)	0.340	0.103
Hou (2013) [2]	China	Discordant twin study	Mixed sex	Parental Hostility	13.86	Conduct Problems	15.36	mixed/ mixed	0.000	0.054
Hou (2013) [4]	China	Discordant twin study	Mixed sex	Parental Hostility	13.86	Conduct Problems	15.36	mixed/ mixed	0.020	0.054
Klahr, McGue (2011)	USA	Adoption study	Mixed sex	Parent-Child Conflict	14.10	Conduct Problems	18.20	caregiver/ self	0.120	0.036
Klahr, Rueter (2011) [1]	USA	Adoption study	Mixed sex	Parent-Child Conflict	14.00	Other DBD	14.00	caregiver/ self	0.400	0.155
Klahr, Rueter (2011) [2]	USA	Adoption study	Mixed sex	Coercive Parenting	14.00	Other DBD	14.00	observer/ self	0.090	0.050
Klahr, Rueter (2011) [3]	USA	Adoption study	Mixed sex	Parent-Child Conflict	14.00	Other DBD	14.00	caregiver/ observer	0.260	0.101
Klahr, Rueter (2011) [4]	USA	Adoption study	Mixed sex	Coercive Parenting	14.00	Other DBD	14.00	observer/ observer	0.330	0.128
Latham (2017) [1]	UK	Discordant sibling study	Mixed sex	Coercive Parenting	3.92	Conduct Problems	5.92	caregiver/ caregiver	0.020	0.160
Latham (2017) [2]	UK	Discordant sibling study	Mixed sex	Coercive Parenting	3.90	Conduct Problems	5.92	caregiver/ caregiver	-0.220	0.150
Lipscomb (2014)	USA	Adoption study	Mixed sex	Overreactive Parenting	3.00	Externalising Behaviour	6.00	caregiver/ caregiver	0.130	0.039

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Long (2015) [1]	USA	Discordant twin study	Mixed sex	Low Parental Warmth	36.69	Conduct Disorder	36.69	self/self	0.037	0.030
Long (2015) [10]	USA	Discordant twin study	Mixed sex	Low Parental Warmth	36.69	Antisocial Personality Disorder	36.69	self/self	0.034	0.030
Long (2015) [11]	USA	Discordant twin study	Mixed sex	Overprotective Parenting	36.69	Antisocial Personality Disorder	36.69	self/self	-0.071	0.030
Long (2015) [12]	USA	Discordant twin study	Mixed sex	Harsh Discipline	36.69	Antisocial Personality Disorder	36.69	self/self	0.015	0.040
Long (2015) [2]	USA	Discordant twin study	Mixed sex	Overprotective Parenting	36.69	Conduct Disorder	36.69	self/self	-0.024	0.030
Long (2015) [3]	USA	Discordant twin study	Mixed sex	Harsh Discipline	36.69	Conduct Disorder	36.69	self/self	0.047	0.040
Long (2015) [4]	USA	Discordant twin study	Mixed sex	Low Parental Warmth	36.69	Conduct Disorder	36.69	self/self	0.025	0.030
Long (2015) [5]	USA	Discordant twin study	Mixed sex	Overprotective Parenting	36.69	Conduct Disorder	36.69	self/self	-0.005	0.030
Long (2015) [6]	USA	Discordant twin study	Mixed sex	Harsh Discipline	36.69	Conduct Disorder	36.69	self/self	0.084	0.040
Long (2015) [7]	USA	Discordant twin study	Mixed sex	Low Parental Warmth	36.69	Antisocial Personality Disorder	36.69	self/self	-0.006	0.030
Long (2015) [8]	USA	Discordant twin study	Mixed sex	Overprotective Parenting	36.69	Antisocial Personality Disorder	36.69	self/self	-0.075	0.030

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Long (2015) [9]	USA	Discordant twin study	Mixed sex	Harsh Discipline	36.69	Antisocial Personality Disorder	36.69	self/self	0.070	0.040
Lysenko (2013) [1]	UK	Discordant sibling study	Males	Harsh Discipline	4.00	Conduct Problems	7.00	caregiver/ caregiver	0.321	0.048
Lysenko (2013) [2]	UK	Discordant sibling study	Females	Harsh Discipline	4.00	Conduct Problems	7.00	caregiver/ caregiver	0.357	0.054
Marceau (2013)	USA	Adoption study	Mixed sex	Overreactive Parenting	1.50	Externalising Behaviour	2.25	caregiver/ caregiver	0.150	0.076
Mark (2017) [2]	UK	Discordant sibling study	Mixed sex	Parent-Child Conflict	10.94	Externalising Behaviour	10.94	self/ caregiver	0.299	0.112
Narusyte (2011) [1]	Sweden	Extended children of twins study	Mixed sex	Parental Criticism	16.20	Externalising Behaviour	16.20	caregiver/ self	0.240	0.038
Narusyte (2011) [2]	Sweden	Extended children of twins study	Mixed sex	Parental Criticism	16.20	Externalising Behaviour	16.20	caregiver/ self	-0.090	0.046
Oliver (2015)	UK	Discordant twin study	Mixed sex	Parental Feeling	5.56	Conduct Problems	9.02	caregiver/ caregiver	0.100	0.039
Pike (1996) [1]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/self	0.330	0.168
Pike (1996) [10]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ caregiver	-0.070	0.104
Pike (1996) [11]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ caregiver	0.100	0.104
Pike (1996) [12]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ observer	0.290	0.148

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Pike (1996) [13]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/self	0.390	0.198
Pike (1996) [14]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/ caregiver	-0.040	0.105
Pike (1996) [16]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/ caregiver	0.100	0.104
Pike (1996) [17]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/ observer	-0.150	0.102
Pike (1996) [18]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver/ self	0.140	0.103
Pike (1996) [19]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver (F)/caregiver (M)	0.190	0.101
Pike (1996) [2]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/ caregiver	-0.130	0.103
Pike (1996) [20]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver/ caregiver	0.540	0.275
Pike (1996) [21]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver/ observer	0.220	0.112
Pike (1996) [22]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ self	-0.070	0.104
Pike (1996) [23]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ caregiver	0.020	0.105
Pike (1996) [24]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ caregiver	0.060	0.104
Pike (1996) [25]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ observer	0.290	0.148

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Pike (1996) [3]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/ caregiver	0.060	0.104
Pike (1996) [4]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	self/ observer	-0.140	0.103
Pike (1996) [5]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver/ self	0.200	0.101
Pike (1996) [6]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver/ caregiver	0.270	0.137
Pike (1996) [7]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver (M)/ caregiver (F)	0.470	0.239
Pike (1996) [8]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	caregiver/ observer	0.290	0.148
Pike (1996) [9]	USA	Discordant twin study	Mixed sex	Parent-Child Conflict	13.71	Other DBD	13.71	observer/ self	-0.110	0.104
Reuben (2016) [2]	USA	Adoption study	Mixed sex	Overreactive Parenting	2.25	Externalising Behaviour	6.50	caregiver/ teacher	0.019	0.068
Reuben (2016) [4]	USA	Adoption study	Mixed sex	Overreactive Parenting	2.25	Externalising Behaviour	6.50	caregiver/ teacher	-0.010	0.072
Richmond (2006) [1]	Not reported	Discordant sibling study	Mixed sex	Parental Hostility	15.00	Externalising Behaviour	15.00	caregiver/ caregiver	0.021	0.005
Richmond (2006) [2]	Not reported	Discordant sibling study	Mixed sex	Parental Hostility	15.00	Externalising Behaviour	15.00	caregiver/ caregiver	0.016	0.005
Richmond (2009) [1]	Not reported	Discordant sibling study	Mixed sex	Parental Hostility	15.00	Externalising Behaviour	15.00	caregiver/ self	0.299	0.100
Richmond (2009) [2]	Not reported	Discordant sibling study	Mixed sex	Parental Hostility	15.00	Externalising Behaviour	15.00	caregiver/ self	0.355	0.118

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{OUT}	Informant	ES	SE
Riggins-Caspers (2003) [1]	Not reported	Adoption study	Mixed sex	Physical Discipline		Conduct Disorder		caregiver/ self	0.190	0.097
Riggins-Caspers (2003) [2]	Not reported	Adoption study	Mixed sex	Harsh Discipline		Conduct Disorder		caregiver/ self	0.250	0.097
Riggins-Caspers (2003) [3]	Not reported	Adoption study	Mixed sex	Physical Discipline		Oppositional Defiant Disorder		caregiver/ self	0.380	0.115
Riggins-Caspers (2003) [4]	Not reported	Adoption study	Mixed sex	Harsh Discipline		Oppositional Defiant Disorder		caregiver/ self	0.420	0.127
Rolon-Arroyo (2018) [1]	USA	Within-person fixed effects	Mixed sex	Overreactive Parenting	5.76	Conduct Disorder	6.74	caregiver/ caregiver	0.110	0.050
Rolon-Arroyo (2018) [2]	USA	Within-person fixed effects	Mixed sex	Overreactive Parenting	5.76	Oppositional Defiant Disorder	6.74	caregiver/ caregiver	0.180	0.050
Roos (2016)	USA	Adoption study	Mixed sex	Parental Involvement	4.50	Externalising Behaviour	6.50	caregiver/ caregiver	0.139	0.066
Samek (2014) [1]	USA	Adoption study	Mixed sex	Parental Involvement	18.24	Antisocial Personality Disorder	23.00	self/self	-0.300	0.153
Samek (2014) [2]	USA	Adoption study	Mixed sex	Parent-Child Conflict	18.24	Antisocial Personality Disorder	23.00	self/self	0.370	0.188
Samek (2014) [3]	USA	Adoption study	Mixed sex	Parent-Child Relationship	18.24	Antisocial Personality Disorder	23.00	self/self	0.270	0.137
Shelton (2008) [1]	UK	Discordant twin study	Mixed sex	Parental Hostility	15.28	Conduct Problems	22.28	self/self	0.279	0.071
Shelton (2008) [2]	UK	Discordant twin study	Mixed sex	Parental Hostility	15.28	Conduct Problems	22.28	caregiver/ caregiver	0.368	0.094

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Shelton (2008) [3]	UK	Discordant twin study	Mixed sex	Parental Warmth	15.28	Conduct Problems	22.28	self/self	0.031	0.066
Shelton (2008) [4]	UK	Discordant twin study	Mixed sex	Parental Warmth	15.28	Conduct Problems	22.28	caregiver/ caregiver	0.179	0.069
Viding (2009) [1]	UK	Discordant twin study	Mixed sex	Harsh Discipline	7.00	Conduct Problems	7.00	caregiver/ caregiver	0.460	0.139
Viding (2009) [2]	UK	Discordant twin study	Mixed sex	Harsh Discipline	7.00	Conduct Problems	7.00	caregiver/ teacher	0.120	0.036
Viding (2009) [3]	UK	Discordant twin study	Mixed sex	Harsh Discipline	7.00	Conduct Problems	12.00	caregiver/ caregiver	0.200	0.060
Viding (2009) [4]	UK	Discordant twin study	Mixed sex	Harsh Discipline	7.00	Conduct Problems	12.00	caregiver/ teacher	0.070	0.027
Viding (2009) [5]	UK	Discordant twin study	Mixed sex	Harsh Discipline	7.00	Conduct Problems	12.00	caregiver/ caregiver	0.120	0.036
Viding (2009) [6]	UK	Discordant twin study	Mixed sex	Harsh Discipline	7.00	Conduct Problems	12.00	caregiver/ teacher	0.070	0.027
Waller, Hyde (2018) [2]	USA	Discordant twin study	Mixed sex	Harsh Discipline	7.80	Other DBD	7.80	mixed/ caregiver	0.137	0.172
Ganiban (2021) [1]	USA	Adoption study	Mixed sex	Parental Involvement	2.25	Externalising Behaviour	7.00	caregiver(M/ F)/caregiver(F/M)	0.024	0.053
Ganiban (2021) [2]	USA	Adoption study	Mixed sex	Overreactive Parenting	2.25	Externalising Behaviour	7.00	caregiver(M/ F)/caregiver(F/M)	0.051	0.053
Cree (2021) [1]	USA	Adoption study	Mixed sex	Overreactive Parenting	1.50	Externalising Behaviour	4.50	caregiver(M) /caregiver(F)	0.224	0.103

APPENDIX C – META-ANALYSIS (CHAPTER 4)

Reference	Country	QE method	Sex	Exposure	Age^{EXP}	Outcome	Age^{O}_{UT}	Informant	ES	SE
Shewark (2021) [1]	USA	Adoption study	Mixed	Parental Hostility	6.00	Externalising Behaviour	7.00	caregiver/ caregiver	0.420	0.035
Shewark (2021) [3]	USA	Adoption study	Mixed	Parental Hostility	6.00	Externalising Behaviour	7.00	caregiver/ caregiver	0.230	0.040

Table 6. Descriptive summary of the participant characteristics and study features of the studies included in the meta-analysis by risk of bias category.

	High quality (*n* = 78)	**High risk** (*n* = 24)	**Very high risk** (*n* = 57)
Percentage females	48%	51%	48 %
Missing	*1*	*0*	*32*
Percentage mothers	77%	95%	55%
Missing	*23*	*4*	*13*
Percentage fathers	23%	5%	45%
Missing	*23*	*4*	*13*
Majority ethnicity			
African American	0 (0%)	6 (43%)	0 (0%)
Asian	8 (26%)	4 (29%)	0 (0%)
White	19 (61%)	4 (29%)	48 (100%)
Hispanic	4 (13%)	0 (0%)	0 (0%)
Missing	*47*	*10*	*9*
Year of publication			
1996 - 2001	0 (0%)	0 (0%)	24 (42%)
2002 - 2006	20 (26%)	8 (33%)	9 (16%)
2007 - 2011	23 (29%)	4 (17%)	6 (11%)
2012 - 2016	24 (31%)	10 (42%)	14 (25%)
2017 - 2021	11 (14%)	2 (8.3%)	4 (7.0%)
Country			
Canada	0 (0%)	8 (33%)	0 (0%)
China	4 (5.1%)	0 (0%)	0 (0%)
Sweden	2 (2.6%)	0 (0%)	0 (0%)
UK	37 (47%)	2 (8.3%)	4 (8.2%)
UK/USA	0 (0%)	0 (0%)	2 (4.1%)
USA	31 (40%)	14 (58%)	43 (88%)
USA/Puerto Rico	4 (5.1%)	0 (0%)	0 (0%)
Missing	*0*	*0*	*8*
Cohort name			
Beijing Twin Study (BeTwiSt)	4 (5.1%)	0 (0%)	0 (0%)
Boricua Youth Study (BYS; Females)	2 (2.6%)	0 (0%)	0 (0%)
Boricua Youth Study (BYS; Males)	2 (2.6%)	0 (0%)	0 (0%)
Cardiff IVF study (CardiffIVF)	0 (0%)	0 (0%)	2 (4.1%)
Cardiff Study of All Wales and North West of England Twins (CaStANET)	4 (5.1%)	0 (0%)	0 (0%)
Early Growth and Development Study (EGDS)	18 (23%)	0 (0%)	0 (0%)
Environmental Risk (E-Risk) Longitudinal Twin Study	8 (10%)	0 (0%)	0 (0%)
Fragile Families and Child Wellbeing Study (FFCWS)	0 (0%)	6 (27%)	0 (0%)
Healthy Babies Healthy Children (HBHC)	0 (0%)	2 (9.1%)	0 (0%)
National Longitudinal Study of Youth (NLSY79)	0 (0%)	2 (9.1%)	0 (0%)

	High quality (*n* = 78)	**High risk** (*n* = 24)	**Very high risk** (*n* = 57)
National Longitudinal Survey of Children and Youth (NLSCY)	0 (0%)	4 (18%)	0 (0%)
Nonshared Environment and Adolescent Development (NEAD) project	0 (0%)	0 (0%)	24 (49%)
Northeast-Northwest Collaborative Adoption Projects (N2CAP)	0 (0%)	0 (0%)	5 (10%)
Ontario Child Health Study (OCHS)	0 (0%)	2 (9.1%)	0 (0%)
Panel Study of Income Dynamics (PSID) - Child Development Supplement (CDS) study	6 (7.7%)	0 (0%)	0 (0%)
Pittsburgh Youth Study (PYS)	3 (3.8%)	0 (0%)	0 (0%)
Sibling Interaction and Behavior Study (SIBS)	4 (5.1%)	4 (18%)	0 (0%)
Sisters and Brothers Study (SBS)	0 (0%)	0 (0%)	2 (4.1%)
The Twins, Family and Behaviour (TFaB)	2 (2.6%)	0 (0%)	0 (0%)
Swedish Twin Registry	2 (2.6%)	0 (0%)	0 (0%)
Twin Study of Behavioral and Emotional Development in Children (TBED-C) within The Michigan State University Twin Registry (MSUTR)	0 (0%)	0 (0%)	2 (4.1%)
Twins Early Development Study (TEDS)	21 (27%)	2 (9.1%)	2 (4.1%)
Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD)	0 (0%)	0 (0%)	12 (24%)
Wales Adoption Cohort Study (WACS)	2 (2.6%)	0 (0%)	0 (0%)
Missing	*0*	*2*	*8*
Causal inference method			
Adoption study	24 (31%)	4 (17%)	9 (16%)
Discordant sibling study	8 (10%)	18 (75%)	6 (11%)
Discordant twin study	37 (47%)	0 (0%)	40 (70%)
Extended children of twins study	2 (2.6%)	0 (0%)	0 (0%)
IVF study	0 (0%)	0 (0%)	2 (3.5%)
Propensity score matching	4 (5.1%)	0 (0%)	0 (0%)
Within-person fixed effects	3 (3.8%)	2 (8.3%)	0 (0%)
Longitudinal			
Cross-sectional	6 (7.7%)	14 (58%)	46 (81%)
Longitudinal	72 (92%)	10 (42%)	11 (19%)

Note: very high-risk studies scored below 5.5, high-risk scored between 5.5 and 7 and high-quality studies scored above 7 on the adapted Newcastle Ottawa scale. Abbreviations: k = number of studies; % = percentage.

Methods 1. Population Attributable Impact of Negative Parenting on Disruptive Behaviour Disorders

```
library(effectsize)
library(ggplot2)
```

Estimates

Prevalence of disruptive behaviour disorders

Polanczyk et al (2015) estimated that the global prevalence of any disruptive behaviour was 5.7% which is the equivalent of 113 million school-aged children.

```
prev_before <- 0.057
```

Impact of Parent Management Training (PMT) on parenting practices

Jeong et al (2021) estimated the impact of current parenting interventions on parenting practices: SMD [Cohen's d] = 0.33 (*95% CI*: 0.22 to 0.44, P < 0.001). Our estimate for the effect of parenting practices on disruptive behaviour disorders is in Pearson's r therefore we will need to convert the SMD to Pearson's r later.

```
est_int_smd <- 0.33
```

Causal effect of parenting practices on disruptive behaviour disorder

The estimate of the effect of parenting practices on disruptive behaviour disorders is taken from our meta-analysis of quasi-experimental studies, expressed as a 1SD increase in parenting practices leading to a -0.102SD decrease in disruptive behaviour.

```
est_par_r_1sd <- -0.102
```

Calculations

Using these estimates and assuming that the prevalence estimates were based on a normally distributed standardised disruptive behaviour score (a deviation from this, say a skewed Poisson distribution would leave to an underestimate).

Convert estimate of the impact of interventions on parenting practices

We first need to convert this estimate from Cohen's d to Pearson's r.

```
est_interv_r <- d_to_r(est_int_smd)

est_interv_r

## [1] 0.1627988
```

So, a parenting intervention has an effect of 0.16 on parenting practices.

We now must scale our current causal estimate (0.102) from 1SD to this 0.16SD value so that it is relative.

```
est_parent_r <- est_par_r_1sd*est_interv_r

est_parent_r

## [1] -0.01660548
```

So, we estimate that a 0.16SD increase in parenting practices decreases offspring disruptive behaviour by -0.0166.

Derive z-value

Now let us find the z-value which corresponds to an area under the curve equal to our estimate of the prevalence of disruptive behaviour disorders (5.7%)

```
1-pnorm(1.58) # 0.05705343 (5.7%)

## [1] 0.05705343

z_before <- 1.58
```

Calculate new prevalence

```
z_after <- 1.58-est_parent_r

prev_after <- 1-pnorm(z_after) # 0.05592642 (5.59%)

prev_diff <- prev_before-prev_after


z_after

## [1] 1.596605

prev_after
```

```
## [1] 0.05517684
```

```
prev_diff
```

```
## [1] 0.001823161
```

Calculate the original reference population

The Polanczyk et al (2015) paper estimated that 113 million school-aged children = 5.7%. Therefore, we can calculate their estimate for the total population that they used as a reference.

```
total_pop <- 113000000/0.057
```

```
total_pop
```

```
## [1] 1982456140
```

Population attributable impact of negative parenting

So, if an effective hypothetical intervention exists, we can estimate how many children would no longer exhibit clinically relevant disruptive behaviour disorders.

```
(total_pop*prev_before)-(total_pop*prev_after)
```

```
## [1] 3614337
```

```
total_pop*prev_diff
```

```
## [1] 3614337
```

Appendix D. Mendelian randomisation (Chapter 5)

Chapter 5 is based on an article published in *Scientific Reports*:

Karwatowska, L., Frach, L., Schoeler, T., Tielbeek, J. J., Murray, J., de Geus, E., Viding, E., & Pingault, J.-B. (2023). Resting heart rate and antisocial behaviour: A Mendelian randomisation study. *Scientific Reports*, *13*(1), Article 1. https://doi.org/10.1038/s41598-023-37123-y

A PDF of this article is included below.

Table 1. STROBE-MR checklist of recommended items to address in reports of Mendelian randomisation studies.

Note. the page numbers relate to the manuscript not the thesis Chapter.

Item No.	Section	Checklist item	Page No.	Relevant text from manuscript
1	TITLE and ABSTRACT	Indicate Mendelian randomization (MR) as the study's design in the title and/or the abstract if that is a main purpose of the study	1	Title: "Resting heart rate and antisocial behavior: A Mendelian randomization study."
	INTRODUCTION			
2	Background	Explain the scientific background and rationale for the reported study. What is the exposure? Is a potential causal relationship between exposure and outcome plausible? Justify why MR is a helpful method to address the study question	5	Scientific background and rationale: "Several reviews exist on putative risk factors for ASB, which include environmental and neurobiological factors. Physiological markers are particularly important in elucidating potential mechanisms underlying the development of ASB. Exposure: "Of these [risk factors], resting heart rate (RHR), defined as the number of heart beats per minute while at rest, is

				the most well-studied." Causal relationship: "There have been several meta-analyses conducted on this topic, all suggesting a robust association between these two factors and concluding that RHR is causally related to ASB." Justification of MR: "Other genetically informed methods can be used to triangulate these findings by relying on different types of data and assumptions."
3	Objectives	State specific objectives clearly, including pre-specified causal hypotheses (if any). State that MR is a method that, under specific assumptions, intends to estimate causal effects	7	Objectives subsection
	METHODS			
4	Study design and data sources	Present key elements of the study design early in the article. Consider including a table listing sources of data for all phases of the study. For each data source contributing to the analysis, describe the following:	7	Study design subsection

	a)	Setting: Describe the study design and the underlying population, if possible. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection, when available.	Available in original study	
	b)	Participants: Give the eligibility criteria, and the sources and methods of selection of participants. Report the sample size, and whether any power or sample size calculations were carried out prior to the main analysis	Available in original study	
	c)	Describe measurement, quality control and selection of genetic variants	Available in original study	
	d)	For each exposure, outcome, and other relevant variables, describe methods of assessment and diagnostic criteria for diseases	eTable 1	
	e)	Provide details of ethics committee approval and participant informed consent, if relevant	Available in original study	
5	Assumptions	Explicitly state the three core IV assumptions for the main analysis (relevance, independence, and exclusion restriction) as well assumptions for any additional or sensitivity analysis	eMethods	"To evaluate causal effects, the genetic variants must satisfy the following three instrumental variable assumptions: the genetic variants indexing the exposure must be (1) associated with the exposure (relevance); (2) independent of confounders of the exposure-

				outcome relationship (exchangeability); and (3) only associated with the outcome through the exposure (exclusion restriction)"
6	Statistical methods: main analysis	Describe statistical methods and statistics used		
	a)	Describe how quantitative variables were handled in the analyses (i.e., scale, units, model)	7 & 8 eMethods	Data sources and measures subsection
	b)	Describe how genetic variants were handled in the analyses and, if applicable, how their weights were selected	eMethods	SNP selection subsection
	c)	Describe the MR estimator (e.g. two-stage least squares, Wald ratio) and related statistics. Detail the included covariates and, in case of two-sample MR, whether the same covariate set was used for adjustment in the two samples	eMethods	SNP selection subsection
	d)	Explain how missing data were addressed	NA	NA
	e)	If applicable, indicate how multiple testing was addressed	NA	NA
7	Assessment of assumptions	Describe any methods or prior knowledge used to assess the assumptions or justify their validity	9 eMethods	Sensitivity analyses subsection
8	Sensitivity analyses and additional analyses	Describe any sensitivity analyses or additional analyses performed (e.g. comparison of effect estimates from different approaches, independent replication, bias analytic techniques, validation of instruments, simulations)	9 & 10 eMethods	Sensitivity analyses; Univariable MR Analysis with Heart Rate Variability; Multivariable MR Analysis with Resting Heart Rate and Heart

				Rate Variability; LD Score Regression subsections
9	Software and pre-registration			
	a)	Name statistical software and package(s), including version and settings used	9	Statistical analyses subsection
	b)	State whether the study protocol and details were pre-registered (as well as when and where)	NA	NA
	RESULTS			
10	Descriptive data			
	a)	Report the numbers of individuals at each stage of included studies and reasons for exclusion. Consider use of a flow diagram	NA	NA
	b)	Report summary statistics for phenotypic exposure(s), outcome(s), and other relevant variables (e.g. means, SDs, proportions)	NA	NA
	c)	If the data sources include meta-analyses of previous studies, provide the assessments of heterogeneity across these studies	NA	NA
	d)	For two-sample MR: i. Provide justification of the similarity of the genetic variant-exposure associations between the exposure and outcome samples ii. Provide information on the number of individuals who overlap between the exposure and outcome studies	7	Data sources and measures subsection
11	Main results			
	a)	Report the associations between genetic variant and exposure, and between genetic variant and outcome, preferably on an interpretable scale	NA	MA

	b)	Report MR estimates of the relationship between exposure and outcome, and the measures of uncertainty from the MR analysis, on an interpretable scale, such as odds ratio or relative risk per SD difference	9	Univariable MR analyses with resting heart rate subsection
	c)	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA	NA
	d)	Consider plots to visualize results (e.g. forest plot, scatterplot of associations between genetic variants and outcome versus between genetic variants and exposure)	eFigures 1-3	
12	Assessment of assumptions			
	a)	Report the assessment of the validity of the assumptions	9	Sensitivity analyses subsection
	b)	Report any additional statistics (e.g., assessments of heterogeneity across genetic variants, such as I^2, Q statistic or E-value)	9	Sensitivity analyses subsection
13	Sensitivity analyses and additional analyses			
	a)	Report any sensitivity analyses to assess the robustness of the main results to violations of the assumptions	9	Sensitivity analyses subsection
	b)	Report results from other sensitivity analyses or additional analyses	9	Sensitivity analyses subsection
	c)	Report any assessment of direction of causal relationship (e.g., bidirectional MR)	9	Sensitivity analyses subsection
	d)	When relevant, report and compare with estimates from non-MR analyses	NA	NA

	e)	Consider additional plots to visualize results (e.g., leave-one-out analyses)	NA	NA
	DISCUSSION			
14	Key results	Summarise key results with reference to study objectives	11	First paragraph
15	Limitations	Discuss limitations of the study, taking into account the validity of the IV assumptions, other sources of potential bias, and imprecision. Discuss both direction and magnitude of any potential bias and any efforts to address them	12	Strengths and limitations subsection
16	Interpretation			
	a)	Meaning: Give a cautious overall interpretation of results in the context of their limitations and in comparison, with other studies	11	Paragraphs 2-5
	b)	Mechanism: Discuss underlying biological mechanisms that could drive a potential causal relationship between the investigated exposure and the outcome, and whether the gene-environment equivalence assumption is reasonable. Use causal language carefully, clarifying that IV estimates may provide causal effects only under certain assumptions	NA	NA
	c)	Clinical relevance: Discuss whether the results have clinical or public policy relevance, and to what extent they inform effect sizes of possible interventions	13	Conclusions
17	Generalizability	Discuss the generalizability of the study results (a) to other populations, (b) across other exposure periods/timings, and (c) across other levels of exposure	12	Strengths and limitations subsection
OTHER INFORMATION				

18	Funding	Describe sources of funding and the role of funders in the present study and, if applicable, sources of funding for the databases and original study or studies on which the present study is based	15	Funding subsection
19	Data and data sharing	Provide the data used to perform all analyses or report where and how the data can be accessed and reference these sources in the article. Provide the statistical code needed to reproduce the results in the article, or report whether the code is publicly accessible and if so, where	15	Data and data sharing subsection
20	Conflicts of Interest	All authors should declare all potential conflicts of interest	15	Conflicts of interest subsection

Appendix E. Trajectories of adversity (Chapter 6)

Table 1 The model fit indices from the 2-class, 3-class, 4-class, 5-class, and 6-class group-based multi-trajectory models.

		Group 1				Group 2				Group 3				Group 4			
	BIC	**N_{IND}**	**%**	**AveP**	**OCC**	**N_{IND}**	**%**	**AveP**	**OCC**	**N_{IND}**	**%**	**AveP**	**OCC**	**N_{IND}**	**%**	**AveP**	**OCC**
2	-7588470	1,542,313	76.33	0.97	9.08	478,225	23.67	0.93	40.02	-	-	-	-	-	-	-	-
3	-7499170	1,467,845	72.65	0.95	70.99	275,118	13.62	0.79	23.67	277,575	13.74	0.96	14.39	-	-	-	-
4	-7402121	1,436,334	71.09	0.94	64.70	86,319	4.27	0.78	78.02	227,836	11.28	0.79	29.42	270,049	13.37	0.94	10.41
5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Abbreviations: BIC: Bayesian information criterion; AveP: average posterior probabilities of assignment; OCC: odds of correct classification.

Table 2 Tests of the proportional hazards assumption for death, diagnoses, and crime for the whole sample

Outcome	Chi^2	Df	p
Death			
Trajectory group	8.67	3	0.034
Sex	81.82	1	< 0.0001
Birth year	1.31	1	0.252
GLOBAL	91.60	5	0.000
Diagnoses			
Trajectory group	56.56	3	< 0.0001
Sex	55.99	1	< 0.0001
Birth year	266.20	1	< 0.0001
GLOBAL	347.24	5	< 0.0001
Crime			
Trajectory group	93.52	3	< 0.0001
Sex	242.47	1	< 0.0001
Birth year	144.43	1	< 0.0001
GLOBAL	508.36	5	< 0.0001

Table 3 Tests of the proportional hazards assumption for death, diagnoses, and crime for males only

Outcome	Chi^2	Df	p
Death			
Trajectory group	3.81	3	0.283
Birth year	0.89	1	0.345
GLOBAL	4.50	4	0.343
Diagnoses			
Trajectory group	48.02	3	0.000
Birth year	252.74	1	0.000
GLOBAL	273.54	4	0.000
Crime			
Trajectory group	69.98	3	0.000
Birth year	114.70	1	0.000
GLOBAL	205.40	4	0.000

Table 4 Tests of the proportional hazards assumption for death, diagnoses, and crime for females only

Outcome	Chi^2	Df	p
Death			
Trajectory group	8.36	3	0.039
Birth year	2.16	1	0.142
GLOBAL	10.40	4	0.034
Diagnoses			
Trajectory group	8.77	3	0.032
Birth year	23.02	1	0.000
GLOBAL	27.56	4	0.000
Crime			
Trajectory group	13.53	3	0.004
Birth year	8.62	1	0.003
GLOBAL	26.03	4	0.000

Figure 1 Two-class multi-trajectory models

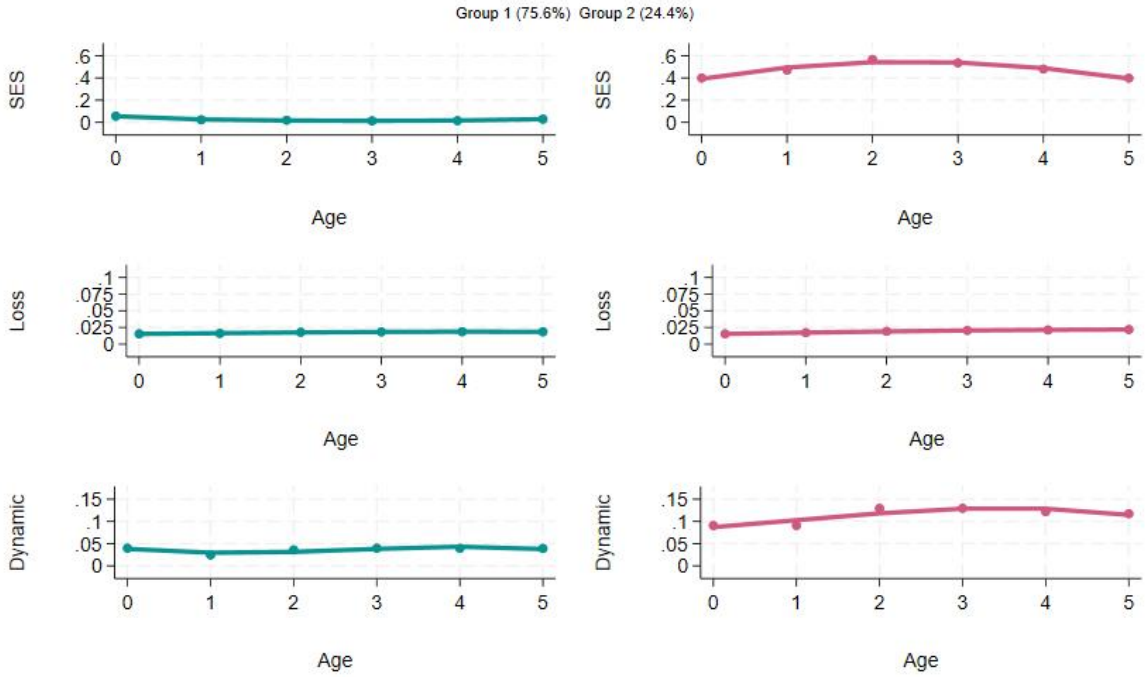


Figure 2 Three-class multi-trajectory models

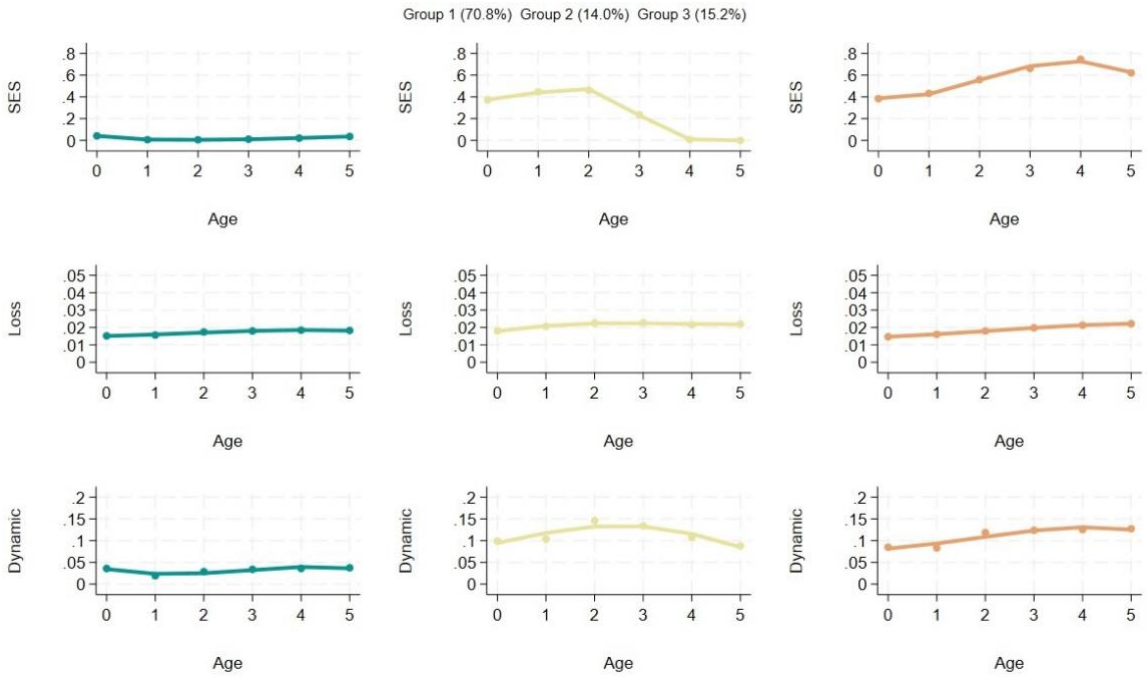


Figure 3 Four-class multi-trajectory models

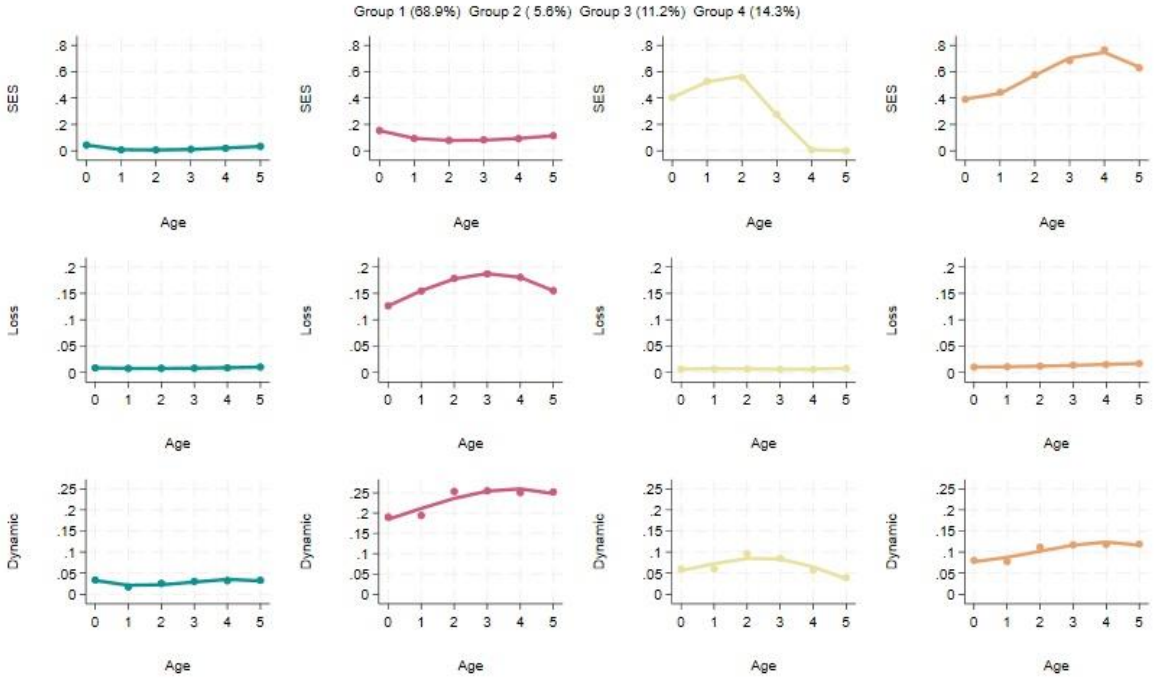


Figure 4 Five-class multi-trajectory models

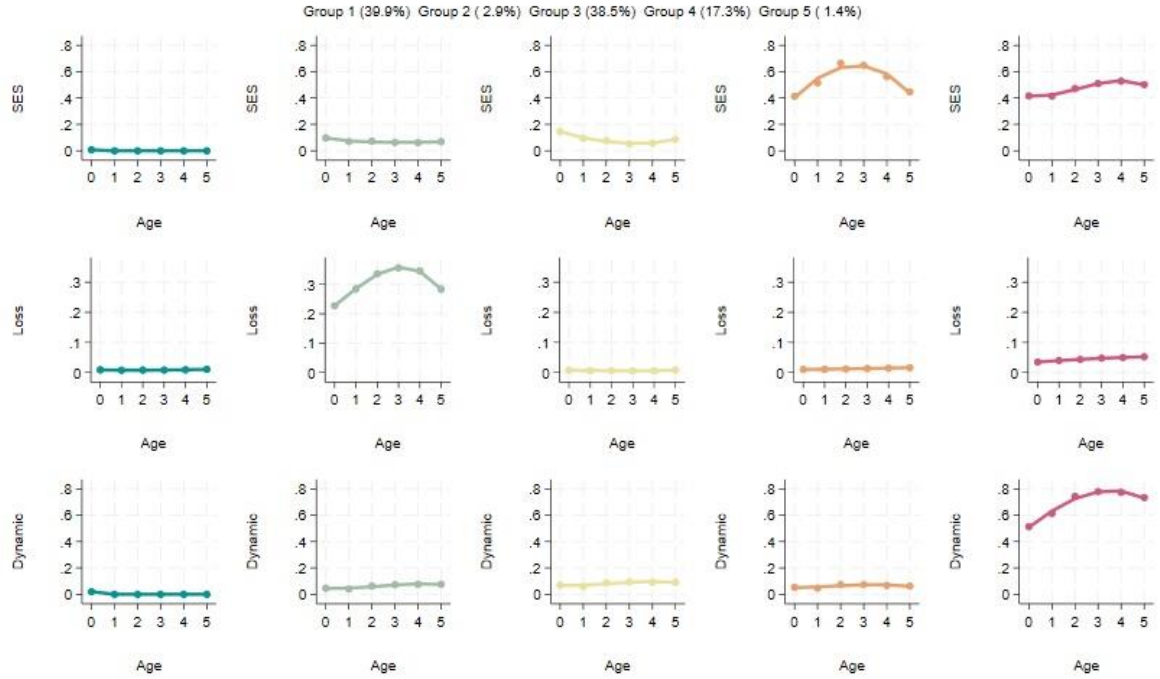


Figure 5 Six-class multi-trajectory models

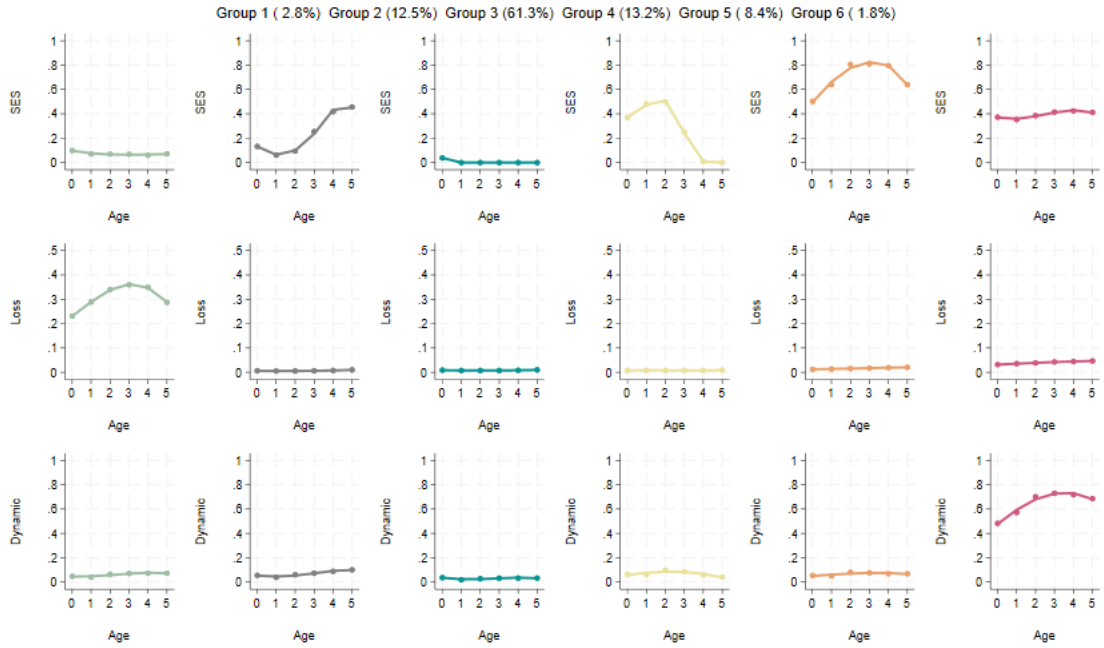


Figure 6 Schoenfeld's residuals – death

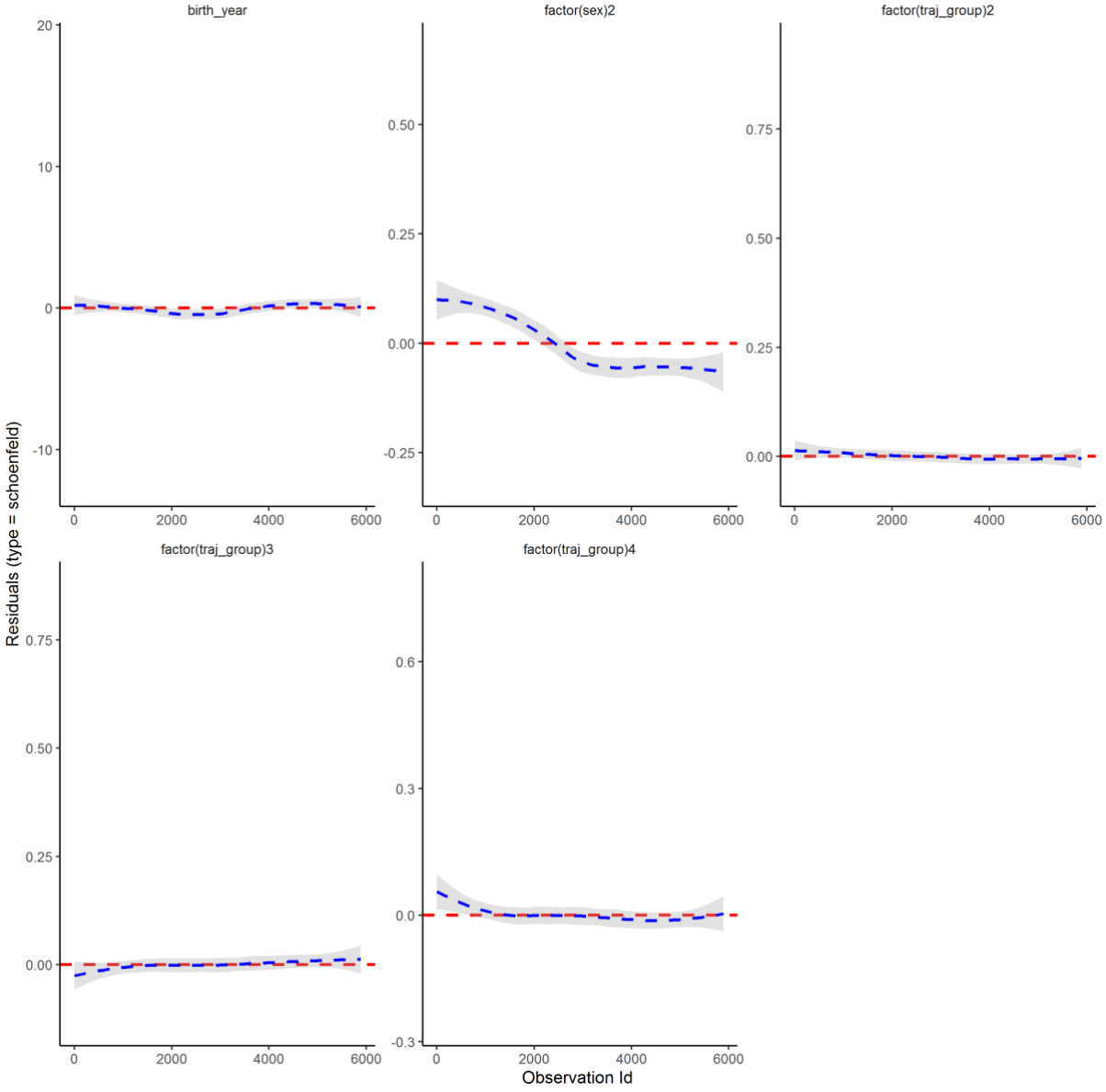


Figure 7 Schoenfeld's residuals –diagnoses

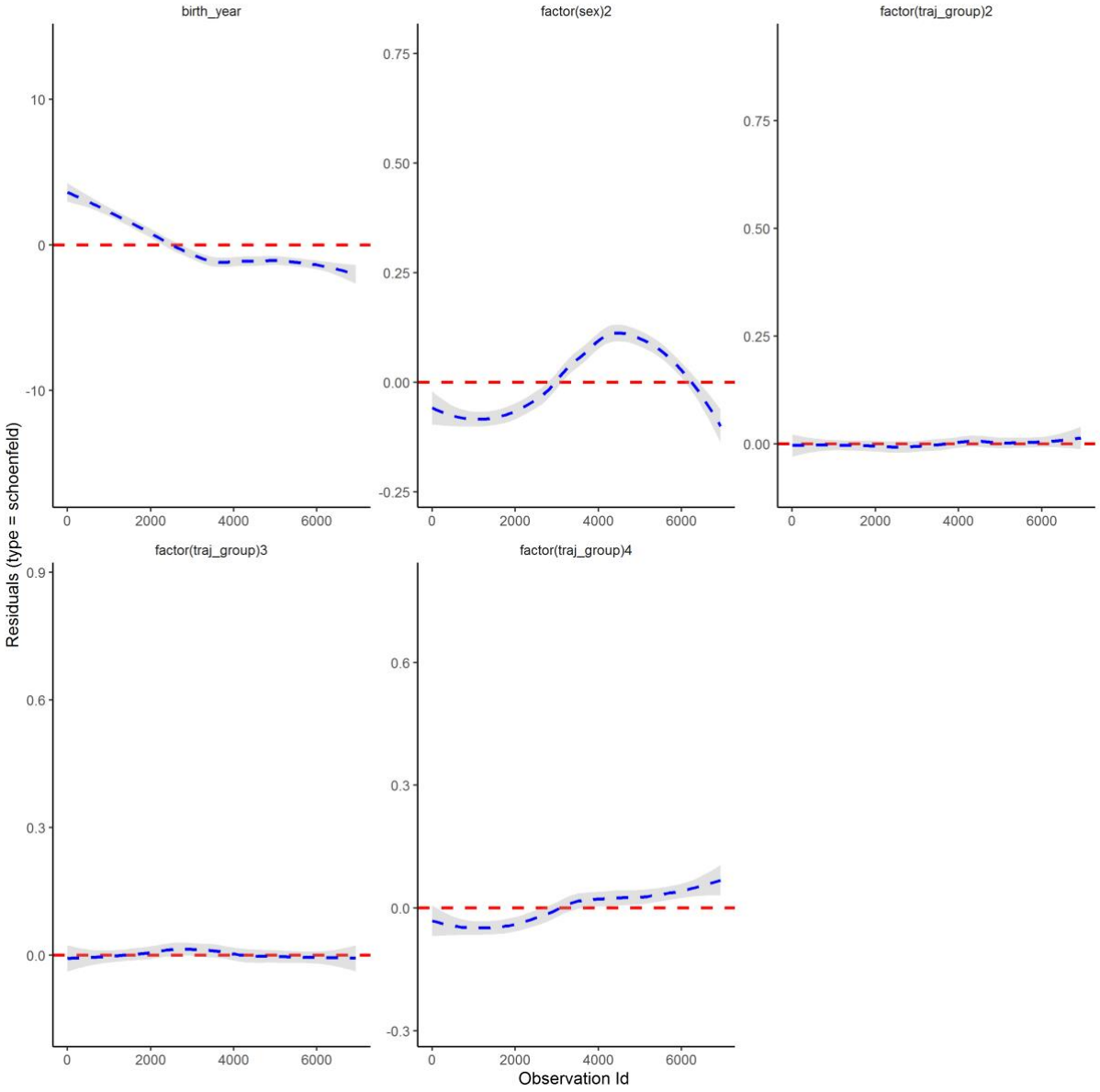


Figure 8 Schoenfeld's residuals – crime