Using Genetically Informative Family Data to Examine the Early Development of Complex Traits

Chloe Austerberry

Submitted for the degree of Doctor of Philosophy

UCL

Declaration

I, Chloe Austerberry, 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.

Signature:

Date: 02.09.2022

Abstract

This thesis triangulates evidence from three genetically sensitive family designs—classical twin studies, a parent-offspring adoption study, and genomic analysis of mother-father-child trios—to explore the roles of genotypes, environments, and interplay between the two, in the development of complex traits. It uses these designs to address two unresolved questions: (1) What are the earliest manifestations of genetic and environmental influences on psychological, developmental, and academic traits? (2) Are the effects of genetic influences on academic outcomes environmentally mediated via evoked differences in the early caregiving environment? These questions were explored in four empirical chapters: The first systematically reviewed and meta-analysed the large twin study literature on genetic and shared and nonshared environmental influences on individual differences in psychological traits and developmental milestones in infancy. The second examined the earliest manifestations of genetic influences on academic abilities, demonstrating that variation in language may be an important early manifestation of genetic influences on later academic performance. The third and fourth chapters examined, using two different genetically informative methods (the adoption design and genomic analysis of mother-father-child trios), whether the effects of genetic influences on academic performance were mediated via evoked differences in early caregiving. One, but not the other, found evidence that parents may adjust their parenting based on their children's education-associated genetic predispositions. Neither found evidence that parenting mediated genetic effects on academic performance, but they did both replicate the finding that language mediates genetic effects on academic performance. Overall, this thesis provides new information about the earliest manifestations of genetic and environmental effects, and interplay between the two, on individual differences in complex traits. In doing so it demonstrates that genetically sensitive family-based research can be leveraged not only to further knowledge of genetic influences but also to better understand environmental pathways in complex trait development.

Table of Contents

List of Tables

List of Figures

List of Tables in Appendices

List of Figures in Appendices

Impact statement

Early childhood represents the most rapid period of postnatal growth and development (Lejarraga, 2012), and research indicates it is a sensitive period for many psychological and developmental traits (Kumsta et al., 2015, Nelson et al., 2019, Rutter 1998). Evidence suggests that complex trait variation can be attributed to a combination of genetic and environmental influences (Polderman et al., 2015). Consequently, to gain understanding of complex trait development in early childhood, it is important to consider both genetic and environmental factors. This thesis aimed to uncover the earliest manifestations of genetic and environmental influences on psychological and developmental traits. It also aimed to uncover environmental mechanisms though which genetic influences might be mediated, focusing particularly on the pathways to intellectual and academic outcomes, which are among the strongest predictors of lifelong success, health, and longevity (Deary et al., 2010; Hummer & Hernandez, 2013; Kosik et al., 2018).

This thesis presents the first comprehensive synthesis of all twin studies of psychological traits and developmental milestones in infancy, offering new insight into the degree to which variation in key domains of infant functioning can be attributed to variation in genes and the shared and nonshared environment. For geneticists, these results provide a metric for how much variation in psychological, developmental, and education-associated traits in infancy and childhood can be attributed to genes as opposed to other factors, offering a guide for future gene discovery research and efforts to uncover the causes of complex trait variation. For clinicians, particularly those working with children, these findings provide an indication of the extent to which family history and environmental factors may predict important outcomes in infancy and childhood, including outcomes that may be early markers of subsequent healthy or pathological development. These findings also have the potential to improve public perceptions on nature and nurture by, for example, dispelling widely held beliefs (which may place undue pressure on parents) that infants are shaped entirely by their environments, or that family history entirely predetermines child health.

This thesis also provides mixed evidence about genetic and environmental influences, and interplay between the two, in the development of academic skills. It uncovered evidence from one study, but not a second, that parents may adjust their parenting based on their children's education-associated genetic predispositions. These mixed findings likely reflect the complexity of the pathways from genes to behaviour and pave the way for future research into gene-environment interplay. The thesis demonstrated more robustly (with replication in studies using two very different research designs) that early language may be a marker of genetic influences on later academic outcomes. Given the apparent importance of intellectual and academic outcomes for lifelong health and wellbeing, this represents a critical finding. For parents, educators, clinicians, and researchers, this points to early language as key aspect of development to focus attention on, research further, and potentially build promotive and preventative interventions around.

Publications and Conference Presentations Associated with Thesis

Publications

Austerberry, C., & Fearon, P. (2021). An overview of developmental behavioral genetics. In L. Provenzi & R. Montirosso (Eds.), *Translational Epigenetics, Developmental Human Behavioral Epigenetics* (pp. 59-80). Academic Press. https://doi.org/10.1016/B978-0-12-819262-7.00003-9.

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M. & Reiss, D. (2022). Early manifestations of intellectual performance: Evidence that genetic effects on later academic test performance are mediated through verbal performance in early childhood. *Child Development*, *93*(2), e188-e206. https://doi.org/10.1111/cdev.13706

Austerberry, C., Mateen, M., Fearon, P., Ronald, A. (2022). Heritability of Psychological Traits and Developmental Milestones in Infancy: A Systematic Review and Meta-Analysis. *JAMA Network Open, 5*(8):e2227887. https://doi.org/10.1001/jamanetworkopen.22.27887

Corfield, E., Frei, O., Shadrin, A. A., Rahman, Z., Lin, A., Athanasiu, L., Akdeniz, B. C., Hannigan, L., Wootton, R., Austerberry, C., Hughes, A., Tesli, M., Westlye, L. T., Njølstad, P., Magnus, P., Davies, N. M., Appadurai, V., Hemani, G., Hovig, E., Zayats, T., Ask, H., Reichborn-Kjennerud, T., Andreassen, O. A. & Havdahl, A. (2022). The Norwegian Mother, Father, and Child cohort study (MoBa) genotyping data resource: MoBaPsychGen pipeline v.1. Preprint: https://www.biorxiv.org/content/10.1101/2022.06.23.496289v2

Conference Presentations

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Reiss, D., Shaw, D. S., Neiderhiser, J. M. (2020). Early Manifestations of Intellectual Ability: Evidence that Genetic Effects on Achievement are Mediated Through Early Childhood Literacy. Behavior Genetics Association 50th Annual Meeting Abstracts. Behavior Genetics, *50*, 441. https://doi.org/10.1007/s10519-020-10018-8

Austerberry, C., Mateen, M., Fearon, P., Ronald, A. Heritability of Infant Traits: Meta-analysis of Twin Studies of Psychological Traits and Developmental Milestones in Infancy. (2021) Behavior Genetics

Association 51st Annual Meeting Abstracts. *Behavior Genetics 51*, 687–762. https://doi.org/10.1007/s10519-021-10087-3

Austerberry, C., Mateen, M., Fearon, P., Ronald, A. Heritability of Infant Traits: Meta-analysis of Twin Studies of Psychological Traits and Developmental Milestones in Infancy. (2020) International Congress of Infant Studies Virtual Congress, July 6-9, 2020.

Contributions to Empirical Studies

Chapter 2

The candidate, Chloe Austerberry (CA), and her supervisors, Angelica Ronald (AR) and Pasco Fearon (PF), designed the study. CA conducted the database searches. CA and DClinPsy student Maria Mateen (MM) extracted the data. CA and PF developed the analytic plan. CA conducted the data analysis and wrote and revised the manuscript. The manuscript was commented on and critically revised by AR and PF.

Chapter 3

CA designed the investigation. CA and PF developed the analytic plan. CA carried out the data analysis and wrote the manuscript. The data were from a pre-existing dataset: The Early Growth and Development Study (EGDS). The proposal for the study and drafts of the manuscript were critically reviewed and commented on by PF, AR, and EGDS principal investigators, David Reiss (DR), Leslie D. Leve (LDL), Jenae Neiderhiser (JMN), and co-investigators, Jody M. Ganiban (JMG), Misaki N. Natsuaki (MNN) and Daniel S. Shaw (DSS).

Chapter 4

CA designed the investigation. CA and PF developed the analytic plan. CA carried out the data analysis and wrote the manuscript. Data were pre-existing, from EGDS. The proposal for the study and drafts of the manuscript were critically reviewed and commented on by PF, AR, DR, LDL, JMN, JMG, MNN, DSS.

Chapter 5

CA designed the investigation. Data were pre-existing, from the Norwegian Mother, Father and Child Cohort Study (MoBa). Access to the data was facilitated and funded by Mona Bekkhus (MB). The genotyping was overseen by Ole Andreassen (OA), Per Minor Magnus (PMM), and Pål Njølstad (PN). The quality control of the genotyped data was carried out by CA in collaboration with colleagues from the University of Oslo, overseen by Elizabeth Corfield (EC) and Alexandra Havdahl (AH), who also oversaw the imputation of the genetic data. Additional post-imputation quality control was carried out by CA, Tetyana Zayats (TZ), who supervised the genetic data analysis, and postdoctoral researcher Dinka Smajlagic (DS). CA constructed the polygenic scores, performed the statistical modelling, and wrote the manuscript. The proposal for the study and drafts of the manuscript were critically reviewed and commented on by TZ, AR, EC, DS, AH, OA, PMM, PN, MB and PF.

Acknowledgements

Firstly, thank you to my wonderful supervisors, Professor Pasco Fearon and Professor Angelica Ronald, for their invaluable advice, guidance, and encouragement. I feel very lucky to have had you both as supervisors.

Thank you to the UKRI ESRC for funding my PhD. Thank you to my collaborators from EGDS and MoBa. I am particularly grateful to my mentor Professor David Reiss for nurturing my interest in this subject and encouraging me to do the PhD. I am also grateful to him and the rest of the EGDS Investigators for their thoughtful feedback on my work and generosity with their time. Thank you to Dr Mona Bekkhus and Dr Tetyana Zayats, both of whom have taught me a huge amount and been extremely supportive of my work. I am indebted to the wider research teams on EGDS and MoBa, and the families participating in both projects. Without their contributions, this project would not have been possible. Nor would it have been possible without the great colleagues and friends I worked with along the way, particularly Dr Sally O'Keeffe, Dr Wikus Barkhuizen, Dr Georgie Donati, Dr Ollie Pain, Laura Havers, Aislinn Bowler, and Dr Anna Gui.

I am grateful to my friends and family for their love and support, especially my mum, Helen—thank you for always being there to chat to (for hours!). Rolo, Julia, Megha, Mickey, Abi, Iona, Nina, and Bertie, thank you for being there to celebrate the highs and support me through the lows of doing a PhD during a pandemic. Finally, Hum, thank you for being by my side throughout, I couldn't have done it without you.

Declarations of Published Work

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)**:**	
a) **Where was the work published?** (e.g. journal name)	Developmental Human Behavioral Epigenetics: Principles, Methods, Evidence and Future Directions
b) **Who published the work?** (e.g. Elsevier/Oxford University Press):	Elsevier
c) **When was the work published?**	01/11/2020
d) **Was the work subject to academic peer review?**	No
e) **Have you retained the copyright for the work?**	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 1b to include in this thesis portions of the publication named as included in 1a.*	

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)**:**		
a) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv):	Please select.	**If yes, which server?** Click or tap here to enter text.
b) **Where is the work intended to be published?** (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.	
c) **List the manuscript's authors in the intended authorship order:**	Click or tap here to enter text.	
d) **Stage of publication**	Please select.	

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4**):**
The candidate (Chloe Austerberry) and her primary supervisor (Pasco Fearon) co-wrote and critically revised the published book chapter.
4. In which chapter(s) of your thesis can this material be found?
Excerpts appear in Chapter 1: General Introduction
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:	Click or tap here to enter text.	**Date:**	24/08/2022
Supervisor/ Senior Author (where appropriate)**:**	Click or tap here to enter text.	**Date:**	24/08/2022

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

6. For a research manuscript that has already been published (if not yet published, please skip to section 2)**:**		
f) Where was the work published? (e.g. journal name)	JAMA Network Open	
g) Who published the work? (e.g. Elsevier/Oxford University Press):	American Medical Association	
h) When was the work published?	22/08/2022	
i) Was the work subject to academic peer review?	Yes	
j) Have you retained the copyright for the work?	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 1b to include in this thesis portions of the publication named as included in 1a.*		
7. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)**:**		
e) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):	Please select.	**If yes, which server?** Click or tap here to enter text.
f) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.	
g) List the manuscript's authors in the intended authorship order:	Click or tap here to enter text.	
h) Stage of publication	Please select.	
8. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)**:**		

The candidate, Chloe Austerberry (CA), and her supervisors, Angelica Ronald (AR) and Pasco Fearon (PF), designed the study. CA conducted the database searches. CA and DClinPsy student Maria Mateen (MM) extracted the data. CA and PF developed the analytic plan. CA conducted the data analysis and wrote and revised the manuscript. The manuscript was commented on and critically revised by AR and PF.
9. In which chapter(s) of your thesis can this material be found?
Chapter 2
10. 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:	Click or tap here to enter text.	**Date:**	24/08/2022
Supervisor/ Senior Author (where appropriate)**:**	Click or tap here to enter text.	**Date:**	24/08/2022

UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

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

k) **Where was the work published?** (e.g. journal name)	Child Development
l) **Who published the work?** (e.g. Elsevier/Oxford University Press):	Wiley
m) **When was the work published?**	16/11/2021
n) **Was the work subject to academic peer review?**	Yes
o) **Have you retained the copyright for the work?**	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 1b to include in this thesis portions of the publication named as included in 1a.*

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

i) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv):	Please select.	**If yes, which server?** Click or tap here to enter text.
j) **Where is the work intended to be published?** (e.g. names of journals that you are planning to submit to)	Click or tap here to enter text.	
k) **List the manuscript's authors in the intended authorship order:**	Click or tap here to enter text.	
l) **Stage of publication**	Please select.	

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

The candidate, Chloe Austerberry (CA), designed the investigation. CA and her primary supervisor, Pasco Fearon (PF) developed the analytic plan. CA carried out the data analysis and wrote the manuscript. The data were from a pre-existing dataset: The Early Growth and Development Study (EGDS). The proposal for the study and drafts of the manuscript were critically reviewed and commented on by PF, Angelica Ronald (the candidate's second supervisor), EGDS principal investigators David Reiss, Leslie D. Leve and Jenae Neiderhiser, and EGDS co-investigators Jody M. Ganiban, Misaki N. Natsuaki and Daniel S. Shaw.
14. In which chapter(s) of your thesis can this material be found?
Chapter 3
15. 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:	Click or tap here to enter text.	**Date:**	24/08/2022
Supervisor/ Senior Author (where appropriate)**:**	Click or tap here to enter text.	**Date:**	24/08/2022

Chapter 1: General Introduction

Excerpts from this chapter were published in *Developmental Human Behavioral Epigenetics Principles, Methods, Evidence, and Future Directions* (Austerberry & Fearon, 2021).

Findings from behavioural genetics research have largely put to rest the somewhat stale nature-versus-nurture debate, by demonstrating through decades of research using a range of genetically informative designs that individual differences in virtually all complex traits and behaviours appear to be influenced by a combination of genes and the environment. Key contemporary questions in behavioural genetics instead concern the timing and interplay of genetic and environmental influences: *when* genetic and environmental factors exert their influence, *how* genes and the environment work together, and *which* genetic variants and environmental influences contribute to individual differences in behaviour. With a particular focus on the development of psychological traits, developmental milestones, and academic outcomes in early and middle-childhood, this thesis aims to address some of these contemporary questions by combining evidence from three genetically sensitive family-based research designs: the classical twin design, the parent-offspring adoption design, and genomic analysis of mother-father-child trios. Specifically, this thesis examines: (1) the earliest manifestations of genetic, shared and nonshared environmental influences on psychological traits, developmental milestones, and academic outcomes, and (2) whether the effects of genetic influences on intellectual and academic outcomes are environmentally mediated via evoked differences in the early caregiving environment.

This first chapter reviews the questions and methods that will be examined in the empirical chapters that follow. It begins by providing a brief overview of behavioural genetics history. Next it outlines core concepts and methods in the field (particularly those used in this thesis), presenting illustrative examples that typify how complex psychological traits and developmental milestones have been studied through the lens of behavioural genetics. Finally, it reviews key findings from behavioural genetics, highlighting the gaps in the literature that will be addressed by the current work, with a particular focus on psychological traits, developmental milestones, and academic outcomes in early and middle childhood.

1.1 A Brief History of Behavioural Genetics

The science of behavioural genetics is thought to have begun in the 19th century with the work of Francis Galton (1822–1911). Influenced by his cousin Charles Darwin's Theory of Natural Selection, Galton was the first to research the inheritance of "mental powers", coined the scientific use of the phrase "nature versus nurture," and first suggested the use of twins and adoptees to study heritability (h^2, which is the proportion of phenotypic variation in a population that can be attributed to genetic differences). However, Galton also became known as the father of the eugenics movement, and it was behavioural genetics' association with this movement that almost entirely discredited it. Following the rediscovery of Mendel's Laws of heredity in 1900, genetically informative research had been on the rise in the early 20th century. A particularly important discovery was that Mendelian laws of single-gene inheritance could be applied to polygenic traits (Fisher, 1918; Wright, 1921), which are produced by multiple-gene effects and are the focus of quantitative genetics research. In the early-to-mid 20th Century the discipline fell out of favour due to its links with the genocide in Nazi Germany and the British and American eugenics movements, and thus environmentalism prevailed for several decades. Behavioural genetics enjoyed a gradual resurgence through the latter half of the 20th century, leading to the more nuanced position held by behavioural scientists today, which acknowledges joint and combined influences of genes and environments on individual differences in complex traits. This resurgence was largely thanks to developments in quantitative genetic methods such as twin and adoption designs. More recently the Human Genome Project and developments in statistical genetics and genomics have added to our understanding of underlying genetic mechanisms and their interplay with the environment. As will be explored below, these newer methods complement but have not supplanted older family-based designs as each rely on different assumptions and provide unique insights.

1.2 Behavioural Genetic Methodology

This section outlines the three methods used in this thesis, which are also some of the most widely used methods in behavioural genetics: twin studies, adoption studies, and research examining associations between behavioural phenotypes and genetic variants.

1.2.1 Twin and Adoption Studies

Twinning and adoption are two naturally occurring phenomena that provide researchers with the opportunity to estimate the degree of variation in a trait that can be attributed to genetic and environmental influences and have been the mainstay of human behavioural genetics for the last 50 or more years. Neither of these methods involve any direct measurement of DNA; instead, they rely on relating patterns of phenotypic resemblance to known familial genetic relationships. How closely phenotypic resemblance appears to mirror patterns of genetic relatedness gives us a clue or estimate of the extent to which genetic factors are influencing individual differences in the phenotype in question. Because of the indirect nature of this way of estimating genetic effects, these quantitative genetic methods are “black boxes” with respect to mechanisms—they estimate the overall contribution of genes to variance in a trait but say nothing about the specific genes or genetic mechanisms involved. As outlined below, the intermediate processes acting across development that eventually give rise to a relationship between a psychological phenotype and the genome are likely extremely complex.

Twin Studies. The classical twin design is based on the comparison of phenotypic similarity within identical (also known as monozygotic, MZ) twins and fraternal (also known as dizygotic, DZ) twins. The logic is that, as MZ twins share 100% of their genes and DZ twins share on average 50% of their segregating genes, a higher degree of phenotypic similarity within MZ twins compared to within DZ twins indicates genetic influence. The twin design also provides a powerful way of estimating the “pure” effect of the environment because differences between MZ twins can only be due to environmental factors. This logic is applied to statistical models that use MZ and DZ twin correlations to produce what are known as ACE estimates, which partition the phenotypic variance in a population into the following sources: additive genetics (A), the shared environment (C), and the nonshared environment (E, which also includes measurement error). It is important to stress that the two components of environmental influence (C and E) do not reflect specific types of environments in any straightforward sense; they simply describe whether unspecified environments make children in the same family similar (shared environment) or different (nonshared environment). No direct measurement of the environment is involved in making this distinction, and while the shared environment may indeed involve family influences (e.g., parenting), such influences may also be experienced quite differently by siblings within the same family and so could be estimated as nonshared environment as well. A common mistake is to assume that evidence of shared environment implies family influence (when in fact it could reflect the action of any common exposure, including in utero biological exposures) or that evidence of the nonshared environment rules out the role of the family. Additional evidence—such as direct measurement of these candidate mechanisms—is required to rule on these sorts of hypotheses.

An important assumption of the twin method is the equal environments assumption, which asserts that environments are as similar for MZ twins as they are for DZ twins. If this assumption is violated due to environments being more similar for MZ than DZ twins, then estimates of genetic influence will be overestimated. A less commonly appreciated corollary is that DZ twins should not experience more dissimilar environments, which would also tend to inflate estimates of genetic influence. When researchers have tested the equal environments assumption, they have tended to find that it is not violated (Bouchard & Propping, 1993; Derks et al., 2006). Broadly speaking, postnatal environments appear to be just as similar for DZ twins as they are for MZ twins, bolstering confidence in findings from twin studies. The twin method also assumes no assortative mating (which occurs when people select mates who are similar or dissimilar to themselves) because, in the case of positive assortative mating (selecting a more similar mate), the DZ genetic correlation is higher than the 0.50 assumed by the statistical models, which will then underestimate genetic influences. There is robust evidence of positive assortative mating in many complex traits, including height and educational attainment (Robinson et al., 2017), suggesting that twin heritability may be systematically underestimated in these domains.

Twin studies have been by far the most widely used tool in the field of behavioural genetics and have been instrumental in convincingly demonstrating the pervasive influence of genetics on human cognition, personality, and psychopathology (Bouchard & Loehlin, 2001; Polderman et al., 2015). Although this has been a critical achievement of behavioural genetics, contemporary research has generally moved on from simply estimating heritability, to focusing on more complex mechanistic and developmental questions. These include investigating the dynamic role of genetic influences across development, the role of genetics in linking different traits together, and the combined influence of genes and environments and interplay between the two across development. A key early insight in the field was that aspects of the environment, as long as they can be meaningfully measured separately for each twin of a twin pair, can be just as easily subjected to quantitative genetic analysis as measures of behaviour. Doing so makes it possible to observe how genetically influenced characteristics may elicit differences in the environment, reversing the direction of the causal arrow typically proposed by developmental psychologists. The study of so-called gene-environment correlation (*r*GE) has produced a wealth of important findings which challenge simple notions of one-way causation from the environment to development (Plomin & Bergeman, 1991). For instance, more recently, multivariate twin analyses (which estimate genetic and environmental influences on the covariance between traits measured at two or more time points or between two or more phenotypes) have been leveraged to identify environmental mechanisms that might be involved in *r*GE (Tucker-Drob & Harden, 2012), which will be explored in greater detail, below. These findings exemplify a much broader phenomenon: namely, that genes influence the emergence of complex traits through an enormous and complex array of indirect steps (gene transcription, protein synthesis, embryological development, and so on), many of which may involve interplay with the environment.

Finally, twin studies are also capable of identifying gene-environment interaction (GxE), a form of gene-environment interplay in which the strength of genetic influence varies as a function of the environment. Despite the attractiveness of the GxE notion for developmentalists, and the strong evidence of its commonplace contribution to development from animal studies (Cooper & Zubek, 1958; Dick, 2011), demonstrating GxE in human populations has proved difficult. Although there may be several reasons for this, a key issue is that, by their very nature, GxE effects are dependent on the level of the environmental exposure, which may be highly variable from one population to another. GxE effects are also highly dependent on the scaling of the measurements and especially on range limitations, such as floor and ceiling effects (Molenaar & Dolan, 2014). Sophisticated psychometric techniques are being developed to try to address these difficulties (Molenaar & Dolan, 2014), but have not yet been widely adopted.

Adoption Studies. Adoption is a natural experiment that creates “genetic” and “environmental” relatives. “Genetic relatives” are genetically related individuals who do not share the same family environment, e.g., adopted children and their birth parents, or genetically related siblings reared separately. “Environmental relatives” are genetically unrelated individuals who share a common family environment, e.g., adopted children and their adoptive parents, or genetically unrelated siblings raised in the same home. Associations between such relatives on developmentally relevant variables can, under certain assumptions, be used to directly estimate genetic and environmental influences. The adoption design is most suited to estimating genetic and environmental influences if adoptees were placed with their adopted families at or very close to birth, as there is less potential for environmental confounding than in later placed children. Other key threats to the validity of the adoption design for estimating genetic effects are prenatal influences, selective placement, and ongoing contact with birth parents. Notably, while adoption practices at the turn of the 20th century tended to include deliberate selective placement, in recent studies there is either limited evidence of selective placement or detected effects can largely be controlled for (Horn, 1983; Leve, Neiderhiser, et al., 2013b; Rhea et al., 2013).

Although not limited to the adoption design, adoption studies may be particularly vulnerable to the problem of poor representativeness—birth parents and adoptive families may not be representative of the wider population. For example, samples of adoptive families are generally under-represented by those of low socioeconomic status (SES) (Stoolmiller, 1999). As a result, findings from adoption studies may only be generalisable to middle income families. McGue et al. (2007) examined the issue of representativeness in a sample of adoptive and non-adoptive families. They found that adoptive families yielded lower variance in measures of psychopathology and SES than non-adopted families. However, this reduction in variance did not appear to markedly influence estimates of the association between family circumstances and children's outcomes, suggesting that the lack of range in adoptive families' circumstances may not substantially bias inferences drawn from adoption studies about the influence of the environment.

Just like the twin design, the adoption design can use associations between phenotypes (e.g., associations between biological parent phenotypes and child phenotypes), rather than direct measures of DNA, to detect genetic influences. Indeed, in the adoption design, the analysis is generally much simpler than the twin design—the humble correlation, for instance, provides a reasonable estimate of the genetic effect. One obvious and important difference between the twin design and the adoption design is that the latter usually relates two measurements taken at different stages of the lifespan (birth parent in adulthood versus adoptee in childhood), and often using different instruments to do so. This will tend to lead the adoption method to underestimate the true heritability of a trait. In that regard, it is notable, for example, that estimates of genetic influence on antisocial behaviour are considerably lower in adoption studies based on parent-offspring pairs (genetic influence ~30%) compared to twin studies (genetic influence ~45%), whereas sibling-based adoption studies are more consistent with the twin estimates (Rhee & Waldman, 2002). Another important consideration when interpreting adoption studies is that statistical power is often not high, partly because large adoption samples are not easy to obtain and because often only one birth parent is available to provide data (so that only half of the genetic effect is observable).

A key methodological advantage of the adoption design is that it removes a major source of rGE, so-called passive rGE, which arises when biological parents provide a rearing environment that is correlated with their own and their child's genotype. The high level of correlation between genes and environments in biological families can make it difficult to disentangle genetic from environmental influences. In the adoption design, the child's genes become effectively uncorrelated with much of the adoptive family environment, and hence adoption studies are particularly well placed for studying the effects of genes, environments, and interplay between the two. Just like the twin design, the adoption design can be used to study rGE. In the adoption design, active and evocative rGE, which arise when an individual's genetic propensities systematically select or evoke differences in their environment, can be tested by examining associations between birth parent characteristics (which are a proxy for genetic influences on adoptees) and measures of the environment, such as parenting or choice of friendship group.

1.2.2 Genomic Approaches

Heritability estimates from twin and adoption studies reveal nothing about the specific genetic variants that may be influencing individual differences in complex behaviour. The Human Genome Project and emerging genomic technologies have allowed researchers to efficiently genotype human DNA, and test for associations between genetic variants and behavioural phenotypes. Increasingly, behavioural genetics researchers have been attempting to uncover specific genetic mechanisms using these genomic tools. In addition to providing direct clues about biological mechanisms (through the identification of specific genes and their functions) genomic studies also provide a potentially crucial corroboration of the results of twin and adoption studies because they rely on very different assumptions. A major approach in the field is tests of genome-wide association.

Genome-Wide Association Studies. Genome-wide association studies (GWAS) scan very large numbers of commonly occurring genetic variants (single-nucleotide polymorphisms [SNPs]) across the entire genome of many individuals and test for associations between phenotypes and typically millions of SNPs. In effect, GWAS represent a blind (hypothesis-free) search of the genome for evidence of association. Drawing on these observed associations, it is possible to calculate SNP heritability (SNP h^2), which is an estimate of the additive contributions of the measured SNPs to phenotypic variation. As SNP h^2 is limited to the effects of common SNPs, it is expected to be lower than the h^2 estimated by quantitative genetic methods (e.g., twin studies), which capture the effects of all genetic variation (additive and non-additive). Because, on their own, single variants tend to have small effects on complex traits and limited predictive power, statistical techniques (e.g., linkage disequilibrium [LD] score regression and polygenic scores) have been developed that use the summary statistics from GWAS to aggregate the effects of thousands of SNPs simultaneously. For example, polygenic scores are calculated as the weighted sum of phenotype-associated SNPs carried by an individual, to provide an estimate of genetic liability for that phenotype.

On its own, the ability of GWAS to pinpoint causal variants is limited by linkage disequilibrium (LD), which is the phenomenon in which SNPs in close physical proximity to one another are more likely to be inherited together. Consequently, most SNPs identified in GWAS as associated with particular phenotypes actually have no causal association with those phenotypes. Rather, they are correlated with the causal SNPs that they were inherited together with due to LD. As a result, to identify true causal variants, additional steps, such as fine mapping, are needed after GWAS have been conducted.

While GWAS have made important contributions to our understanding of the genetic basis of complex developmental traits, major challenges remain in using these methods for understanding mechanisms of development. The identification of specific genomic variants associated with complex traits gets us a little closer to informative underlying neurobiology, especially when combined with transcriptomics data and modelling based on biological pathways and interactions databases, but there remains a vertiginous gulf between indications of genetic association and mechanistic understanding of development. There are many hurdles to overcome before these genomic techniques can more substantially advance developmental science. One serious barrier for GWAS is the very large sample sizes required, which currently place severe limits on how frequently waves of data can be collected and on the richness of the data that can be captured at that scale. GWAS require very large samples, in part due to the extremely low significance threshold required to account for the many statistical tests being conducted. Realistic and in-depth measurements of the environment, which we know are a crucial part of the picture from quantitative genetic studies, are currently difficult to include in GWAS due to the prodigious costs. Another major limitation of GWAS is that they have been conducted primarily in populations of European descent (Peterson et al., 2019). Some ethnic groups are yet to be sequenced and thus optimal GWAS and genotype imputation for these populations cannot yet be conducted (Tam et al., 2019). PRS currently show poor generalisability in non-European populations and efforts are underway to increase the accuracy of PRS across diverse groups (Wang et al., 2022).

As polygenic risk scores improve, in terms of their generalisability, the range of phenotypic domains captured, and the proportion of variance they explain, it is likely that smaller scale (though still likely requiring participant numbers in the thousands), developmentally richer, longitudinal studies will become more feasible in the coming years. Research using polygenic scores is also amenable, at least in principle, to studying the same kinds of complex developmental questions that quantitative genetics has investigated over the last few decades, including genetic mechanisms of continuity and change, rGE and GxE. For example, when family-based data is combined with polygenic scores, it is possible to examine ‘genetic nurture’, which refers to the effects of parent genes on child phenotypes via mechanisms other than direct genetic transmission (i.e., through environmental pathways). There are two ways in which polygenic scores from families have been used to examine genetic nurture. One is to create two parental polygenic scores: the first, only from alleles that were transmitted from parent to child and, the second, only from alleles that were not transmitted (the latter, referred to as the ‘virtual parent’). The effects on child phenotypes of the ‘virtual parent’ polygenic scores, calculated from non-transmitted alleles, are interpreted as genetic nurture effects (Bates et al., 2018; Kong et al., 2018). Second, genetic nurture can be tested by examining the effects of parent polygenic scores on child phenotypes, while controlling for the child’s polygenic score (Wertz et al., 2020). Similar methods can be used to explore evocative *r*GE, which can be examined by testing associations between child polygenic scores and aspects of the rearing environment (e.g., parenting), while controlling for parent polygenic scores.

As well as highlighting this important application of genomic techniques for studying rGE, literature on genetic nurture highlights one of several significant and not always thoroughly appreciated methodological issues in GWAS-based genetic epidemiology. The vast majority of GWAS do not take account of parental genotype, and, for all but a small number of phenotypes, we currently do not know the degree to which current association estimates are confounded by parental genotype. Genetic epidemiological studies are also quite susceptible to bias due to population stratification—where sub-strata of the population differ both in the prevalence of the phenotype of interest and in their genotypes, leading to artefactual association. Rather than identifying true allele-phenotype associations, spurious associations arise, which are wholly explained by differences in ancestry. Although attempts have been made to statistically control for such stratification using principal components, it is becoming clear that quite subtle ancestral differences may be common and can bias GWAS estimates (Byrne et al., 2020). This, and other forms of bias, continue to be important and active areas of methodological development within the field (Morris et al., 2019).

1.3 Key Interpretative Issues

In outlining the twin and adoption methods above, I already touched on several key interpretative issues that must always be kept in mind when appraising data from quantitative genetics research. One is so critical that it warrants repeating: as black box methods for estimating the overall contribution of heritable genetic factors to complex traits, twin and adoption methods (unless combined with genomic data) say nothing about the underlying biological mechanisms involved. Furthermore, both quantitative genetics methods and methods drawing on genomic data describe the net result of most likely an exceptionally large number of complex gene-environment processes unfolding at multiple levels of biological and social organisation over the course of development. Finding evidence of heritability does not imply simple, unmediated, genetic influence on a trait, and many genetic effects may involve substantial environmental mediation (Rutter, 2000), for example via the mechanisms of rGE that are discussed in further detail, below. Additionally, the estimates of genetic influence that are obtained from quantitative genetics methods and genomics describe the current causes of population differences in a trait, and not the degree to which genetic factors are responsible for a trait in a given individual. Critically, substantial heritability does not imply genetic determinism or immutability. Changes to the environment can lead to mean changes in a population, even for highly heritable phenotypes that remain stable over time. Widely referenced examples of this are physical height and intelligence, both of which have increased substantially over the last century, despite heritability estimates for both remaining high and stable throughout (Fisher, 1918; Flynn, 1987; Lettre, 2011; NCD Risk Factor Collaboration, 2016; Pietschnig & Voracek, 2015). Furthermore, causation does not denote treatment and even phenotypes that are entirely genetic can be changed with purely environmental interventions. The most commonly cited example to illustrate this is phenylketonuria (PKU), which is a genetic condition that leads to the inability to metabolize the amino acid phenylalanine. Untreated, PKU leads to severe damage to the central nervous system, but a comparatively simple environmental intervention—excluding phenylalanine from the diet—entirely prevents any adverse developmental effects, as long as it is introduced shortly after birth. A further, often under-appreciated, interpretative issue concerns the role of GxE. As I noted above, there are significant difficulties in human quantitative genetic studies in properly capturing GxE effects (Dick, 2011), even though most commentators agree that it is highly likely they exist and indeed are prevalent. Ignoring GxE can lead to biases in effect estimates (Eaves & Rao, 1984). As a result, it is helpful to be aware of the consequences of ignored GxE, when appraising studies that report genetic “main effects.” In general, in standard modelling, such as that used in twin analyses, ignored gene-by-common environment interactions will be estimated as genetic effects, whereas ignored gene-by-nonshared environment interactions will be estimated as nonshared environment effects.

1.4 Key Findings from Behaviour Genetics

As twin, adoption, and genomic research designs each depend on different assumptions, confidence in their findings is strengthened when they converge on consistent results. Below I outline some key discoveries in developmental behavioural genetics that are supported by converging evidence from studies using different methods. With a particular emphasis on the development of psychological traits and developmental milestones in early and middle childhood, I also note important gaps in the literature, including those that are addressed by the present thesis.

1.4.1 Heritability of Complex Traits

Behavioural genetics research consistently demonstrates that virtually all psychological and behavioural traits are under genetic influence—what has come to be known as the ‘first law’ of behavioural genetics (Turkheimer, 2000)—but none are entirely heritable. For example, twin and other family-based designs converge on a heritability estimate of around 90% for autism (Sandin et al., 2014; Tick et al., 2016), 74% for ADHD (Faraone & Larsson, 2019), and 50–80% for general cognitive ability (Bouchard & McGue, 1981; C. M. A. Haworth et al., 2010). In a meta-analysis of almost all published twin studies of complex traits (predominantly psychiatric, metabolic and cognitive), the heritability estimate when all traits and age groups were combined was 49% (Polderman et al., 2015), providing compelling evidence for the importance of both genes and environments in complex trait development. We know that heritability changes across the lifespan (Bergen et al., 2007). In spite of this, and despite infancy being a rapid and sensitive period of development that deserves special focus (Bornstein, 2014; Kumsta et al., 2015), pooled estimates of heritability, shared and nonshared environment in infancy have never been calculated. This provides a strong rationale for **Chapter 2** of this thesis, which meta-analyses all twin studies on psychologically-relevant traits and milestones in infants (aged 0–2 years).

Consistent with evidence of heritability from twin and family-based studies, in the last 5–15 years, large-scale GWAS have successfully identified a large number of seemingly reliable (replicable) genotype-phenotype associations for many psychiatric disorders and complex traits (Smoller et al., 2019; Visscher et al., 2017). A surprising finding has been that SNP h^2 and polygenic scores constructed from GWAS summary statistics explain a fraction of the heritability identified in twin and adoption studies, limiting the predictive value of GWAS. This discrepancy has come to be known as ‘missing heritability’. For example, a polygenic score created by summing together the number of outcome-related SNPs from a recent GWAS of total years of education (EduYears), involving over 3 million individuals, explained 12–16% of the variance in educational attainment in independent prediction samples (Okbay et al., 2022). In contrast, a meta-analysis of twin studies on educational attainment arrived at a pooled heritability estimate of 43%—approximately three times higher (Silventoinen et al., 2020). While some of the missing heritability may be explained by rare and ultra-rare variants of large effect, which may never be possible to detect using GWAS, the hope is that with improved methods (such as whole-genome sequencing) and increasing sample sizes, GWAS may soon account for a larger fraction of the heritability of behavioural phenotypes. Multivariate analyses incorporating multiple polygenic scores (Plomin & von Stumm, 2018) and the study of gene-gene and gene-environment interplay are also likely to help explain more of the overall phenotypic variation. Until the heritability gap is reduced, it remains important to triangulate estimates inferred from patterns of family resemblance (e.g., twin and adoption studies) with those derived from polygenic score analyses, particularly as each rely on different assumptions (some of which, as I have discussed, may overestimate heritability and others underestimate it, if violated). The former approach (used in **Chapters 2, 3 and 4**) estimates the overall contribution of genetic differences to individual differences in a trait but provides no information about the specific genes or genetic mechanisms involved. The latter approach (which I use in **Chapter 5**, alongside family data) is more informative about underlying mechanisms but estimates only the additive effects of common genetic variation, thus intrinsically underestimating heritability.

1.4.2 Genetic Influence on the Longitudinal Stability of Complex Traits

Broadly speaking, genetic differences appear to be a more systematic source of influence on variation in complex traits than environmental differences (Plomin, 2018) and longitudinal evidence consistently suggests that the stability of traits across development is largely driven by genetic stability. For example, a meta-analysis of longitudinal twin and adoption studies found that stability in IQ across the lifespan was almost entirely due to common genetic influences (Briley & Tucker-Drob, 2013). Similar findings have been replicated in relation to many other psychologically-relevant traits (Plomin et al., 2016). Given the ubiquity and stability of genetic influences on complex traits, it is important, for purposes of prediction and intervention, to identify genetic influences early on in development. **Chapter 2** aimed to quantify heritability in infancy among all previously examined psychological phenotypes and developmental milestones. **Chapter 3** aimed to identify the earliest manifestations of genetic effects on academic performance, a phenotype with moderate-to-high heritability in adulthood (Silventoinen et al., 2020), which robustly predicts important life outcomes (Hummer & Hernandez, 2013; Kosik et al., 2018). Considering the increasing evidence suggesting the importance of *r*GE, it is important to bear in mind that evidence of genetic influences on stability does not imply that genetically based stability is not in part underpinned by cascading and reinforcing environmental processes that are correlated with genetic differences. **Chapters 4 and 5** of this thesis aimed to test whether genes indirectly effect educational outcomes during early development, via genetically correlated evoked differences in the caregiving environment.

1.4.3 Increasing Heritability of IQ

A counterintuitive but now widely accepted finding is that the heritability of IQ increases across the lifespan, alongside a concurrent reduction of the influence of the environment that is shared by siblings (McGue et al., 1993; Plomin & Deary, 2015). Given the stability of the genome itself, it seems somewhat paradoxical that the heritability of IQ increases across the lifespan. The most plausible explanation appears to be genetic amplification through processes of *r*GE. The idea is that genetic differences can become amplified across development as individuals influence, select and modify environments that are correlated with their genotype (Plomin et al., 1977; Scarr & McCartney, 1983). *r*GE will be discussed in greater detail below and is explored in depth in **Chapters 4 and 5**, both of which test for the presence of *r*GE in cognitive and educational development (the former in an adoption sample and the latter using genome-wide polygenic scores from genetically related mother-father-child trios).

1.4.4 Most Environmental Effects are Nonshared

Developmental psychology has tended to draw attention to the influence of environments that are shared by family members and often assumes, implicitly, that experiences within the family will make siblings similar. However, behavioural genetics has generally found quite limited evidence for shared environmental effects on differences in complex traits within a population. Even if one takes account of the fact that the nonshared environment also captures non-systematic measurement error, the nonshared environment appears to be the primary source of environmental variance for the majority of complex traits (Plomin, 2011; Plomin & Daniels, 1987). However, it would be inaccurate to say that shared environment effects are not important at all. Shared environment effects on cognitive ability may be stronger and more persistent in low-SES populations, although the picture is complex (Tucker-Drob & Bates, 2015; Turkheimer et al., 2003). There is also evidence of shared environmental influences on IQ during childhood (Briley & Tucker-Drob, 2013; C. M. A. Haworth et al., 2010), although by adulthood shared environmental effects on cognitive ability decline essentially to zero (Plomin et al., 2016). The influence of the shared and nonshared environment on variation in complex traits early on in development is currently unclear as the infant twin literature has never been synthesized. **Chapter 2** addresses this gap in a meta-analysis that calculates pooled estimates in infancy, for the first time, of shared and nonshared environmental influences on individual differences in psychological traits and developmental milestones.

1.4.5 Gene-Environment Interplay

As noted above, a major focus of contemporary behaviour genetics research is *how* genes and environments work together to influence phenotypic variation, through the study of GxE and *r*GE. Here I discuss *r*GE findings in greater depth as *r*GE a primary focus of the present thesis.

Gene-Environment Correlation. *r*GE is a central mechanism of interest in the study of development, as it provides important insights into the dynamic interplay between inherited characteristics and the psychosocial environment. Broadly speaking, *r*GE is said to be present when an individual's environment is correlated with their genotype. The near ubiquity of the phenomenon is indicated by a substantial body of behavioural genetics literature. For example, a systematic review of quantitative genetic studies (twin, adoption and step-family designs) examining the heritability of 35 ostensibly environmental measures (including, parenting, family environment, and stressful life events) found a weighted heritability of 27% across each (Kendler & Baker, 2007). This is corroborated by recent evidence of covariation between polygenic variation and environmental exposures such as household income, breastfeeding, and parental age, behaviour and education (Krapohl et al., 2017). Three main forms of *r*GE have been defined in the literature: passive, active, and evocative (Plomin et al., 2016; Scarr & McCartney, 1983).

Passive rGE occurs when the parental genes that are transmitted to the child are also correlated with the rearing environment the parent provides. This makes it difficult to establish whether genotype-phenotype associations are a product of direct genetic transmission, the rearing environment, or a combination of both. Researchers have used a number of genetically informative designs to examine passive *r*GE. For example, using a sample of children conceived via assisted reproduction (either with their parents' eggs and sperm or via sperm, egg, or embryo donation), Rice et al. (2013) found evidence of passive *r*GE when they compared parent-child associations in depressive symptoms within biologically related and non-biologically families.

The fast-growing genomic literature on genetic nurture also highlights the importance of passive *r*GE (Kong et al., 2018; Wang et al., 2021). Implicit in the discovery that the parts of the parental genotype *not* transmitted to offspring predict offspring educational outcomes is the likelihood that the same is true of the parts of the parental genotype that *are* transmitted to offspring—that they influence children not only through direct genetic transmission but also through environmentally mediated pathways. This is consistent with evidence from a study that used UK Biobank data to combine genomic data with the adoption design and found that polygenic scores were twice as predictive of years of education in nonadopted individuals compared with adoptees, suggesting that genetic influences on education are mediated via the home environment (or wider environments associated with the home environment) (Cheesman et al., 2020). Although these findings do not undermine the importance for prediction purposes of identifying genetic associations, they do indicate that it is at least possible that genetic associations occur only because they correlate with unobserved, causal, environmental mechanisms (Koellinger & Harden, 2018). This paves the way for research into environmental mechanisms that correlate with genetic differences (such as our research in **Chapters 4 and 5**) and highlights the importance of controlling for passive *r*GE when investigating the influence of genetic and environmental mechanisms (as was done when I used the adoption design in **Chapters 3 and 4**).

Active rGE occurs when an individual's genetically influenced traits influence the types of environments that they select or choose, such as career or friendship group. There is good reason to believe that active *r*GE becomes more important beyond childhood, once individuals have a greater opportunity for active selection of their environments. For example, a study by Connolly et al. (2015) found that peer pressure encouraging of delinquent behaviour showed limited genetic influence in preadolescence but increasing genetic influence across adolescence, consistent with the notion that active *r*GE starts to become a substantial driver of peer processes during this later period. As the present thesis was focused on early development (infancy and early to middle childhood), rather than adolescence or adulthood (when active *r*GE appears to be most important), active *r*GE was not a subject of our analyses.

Evocative rGE occurs when an individual’s genetically influenced characteristics evoke responses from their environment. For example, in the context of early cognitive development and education performance (which are the focus of **Chapters 3, 4 and 5**), genetic differences may systematically elicit different learning environments from parents and teachers. Recently, genomic methods have been combined with family data to examine evocative *r*GE. For example, (Wertz et al., 2020) found that children’s EduYears polygenic scores predicted the home environment (chaos, safety and tidiness) and maternal parenting (warmth, sensitivity and cognitive stimulation) they received, even after controlling for mothers’ polygenic scores. However, use of genomic data to test for evocative *r*GE is in its infancy. Wider and more comprehensive evidence of the evocative influence of child genes on caregiving behaviour comes from twin and parent-offspring adoption studies. A classic adoption study by Ge et al. (1996) found that children aged 12–18 years whose birth mothers showed high levels of externalizing problems received more negative and harsh parenting from their adoptive parents relative to those children whose birth mothers did not have externalizing problems. Twin and family data have also yielded quite consistent evidence of evocative *r*GE in adolescence. For example, there is evidence from both twin, sibling, and an extended children-of-twins study to suggest that evocative *r*GE may explain the correlation between adolescent externalizing problems and parental negativity (Marceau et al., 2013; Pike et al., 1996). Less attention has been paid to *r*GE in the earliest stages of the lifespan, in spite of the fact that evocative *r*GE may be the most important form of *r*GE in early development, when child effects on caregiving have been well established but active selection of environments is likely to be limited (Bell, 1968). One of the few studies to have examined evocative *r*GE across a wide range of traits in early childhood is the Early Growth and Development Study (EGDS)—a prospective adoption study of early childhood (Leve et al., 2019; Leve, Neiderhiser, et al., 2013b). EGDS has found evidence of evocative effects on parenting in infancy and toddlerhood in relation to genetic risk for internalizing and externalizing psychopathology, low social motivation, and attention-deficit hyperactivity disorder (Elam et al., 2014; Fearon et al., 2015; Harold et al., 2013; Klahr et al., 2017). Data from the EGDS study indicated, for example, that even by 9 months of age, infants of birth mothers with greater externalizing problems evoked more parent negativity in their adoptive mothers, partially accounting for later child behavioural problems at 36 months (Fearon et al., 2015). This tendency was only observed in adoptive families reporting high levels of marital distress, suggesting that relationship factors may buffer or amplify the degree to which the caregiving environment is liable to being evoked by the child's heritable traits (Fearon et al., 2015). This way of understanding *r*GE—with evoked environmental responses potentially playing a causal role in the mechanisms of genetic risk transmission—is likely to be an increasingly important focus for behavioural genetic research in the coming years. Although it has been long hypothesised that evocative *r*GE may account for rising intelligence across the lifespan and across generations (Dickens & Flynn, 2001; Plomin et al., 1977), this hypothesis has never been tested using an adoption design, or genetic data from children and both of their parents. I address these gaps in the literature in **Chapters 4 and 5** of this thesis. The former uses the adoption design to examine evocative *r*GE in language and educational attainment in early to middle childhood. The latter examines the same research questions using polygenic scores from mother-father-child trios.

1.5 Summary and Thesis Aims

The field of behavioural genetics has made vital contributions to developmental science and will continue to be a key framework and set of research tools for psychologists and developmentalists in the coming decades. The reviewed evidence persuasively demonstrates how widespread and ubiquitous genetic factors are in shaping the development of complex traits. It highlights the critical importance of considering the timing of genetic and environmental influences, the dynamic interplay between these two sources of influence, and the possible indirect effects of genes on environments via environmental mechanisms that are correlated with genetic differences. It also underlines how powerful a set of tools family-based genomic and quantitative genetics methods are for studying these key contemporary questions. The present thesis triangulates evidence from three genetically sensitive family-based research designs—the classical twin design, the parent-offspring adoption design, and genomic analysis of mother-father-child trios. It leverages these methods to address key unresolved questions about the earliest manifestations of genetic and environmental influences and gene-environment correlation. These questions were explored in four empirical chapters:

The first two empirical chapters focused on the earliest manifestations of genetic and environmental influences on complex traits. **Chapter 2** meta-analyses the large and comprehensive twin study literature to calculate, for the first time, pooled estimates of genetic, shared, and nonshared environmental influences on individual differences in psychological traits and developmental milestones during infancy—a particularly sensitive and rapid period of development. **Chapter 3** aimed to identify the earliest manifestations of genetic effects on academic performance, a phenotype that powerfully predicts important life outcomes and is moderately to highly heritable in adulthood. Based on wider literature, this chapter considered two possible candidates: early childhood executive function and early language performance. In line with the overarching aim to draw on a variety of genetically informative family-based designs, **Chapter 3** examined the early manifestations of genetic effects using data from a parent-offspring adoption study (to complement the evidence from twin studies in **Chapter 2**).

The final two empirical chapters addressed the question of whether there are indirect genetic effects on academic outcomes via evoked environmental mechanisms that are masked by genotype-environment correlations. **Chapter 4** examined, for the first time using an adoption design (which controls for passive rGE), whether the effects of genetic influences on academic achievement are mediated via evoked differences in early caregiving. Based on results from **Chapter 3**, it also examined whether early language is a mechanism through which caregiving differences are evoked. Driven by the overarching aim to triangulate findings from different research designs, the fourth empirical chapter (**Chapter 5**) addresses the same research questions about rGE as **Chapter 4** but in a different sample (a large birth cohort), using different methods (genomic analysis of mother-father-child trios). This is the first time that evocative rGE in educational attainment has been examined in an analysis using polygenic scores from children and both of their biologically related parents.

The thesis concludes with a general discussion (**Chapter 6**). The discussion summarises and synthesises the findings from the thesis, reflects on its results in the context of wider literature, considers the overall limitations of the four empirical studies, and discusses the implications of the thesis for future research.

1.6 References

Austerberry, C., & Fearon, P. (2021). An overview of developmental behavioral genetics. In L. Provenzi & R. Montirosso (Eds.), *Developmental human behavioral epigenetics* (Vol. 23, pp. 59-80). Academic Press. https://doi.org/10.1016/B978-0-12-819262-7.00003-9

Austerberry, C., Mateen, M., Fearon, P., & Ronald, A. (2022). Heritability of psychological traits and developmental milestones in infancy: A systematic review and meta-analysis. *JAMA Network Open, 5*(8), e2227887-e2227887. https://doi.org/10.1001/jamanetworkopen.2022.27887

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

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75*(2), 81-95. https://doi.org/10.1037/h0025583

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics, 10*(3), 423-433. https://doi.org/10.1375/twin.10.3.423

Bornstein, M. H. (2014). Human infancy…and the rest of the lifespan. *Annual review of psychology, 65*, 121-158. https://doi.org/10.1146/annurev-psych-120710-100359

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*(4498), 1055-1059. https://doi.org/10.1126/science.7195071

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

Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics, 31*(3), 243-273. https://doi.org/10.1023/A:1012294324713

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24*(9), 1704-1713. https://doi.org/10.1177/0956797613478618

Byrne, R. P., van Rheenen, W., van den Berg, L. H., Veldink, J. H., McLaughlin, R. L., & Project MinE ALS Gwas Consortium. (2020). Dutch population structure across space, time and GWAS design. *Nature Communications, 11*(1), 4556. https://doi.org/10.1038/s41467-020-18418-4

Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., Eley, T. C., & Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological Science, 31*(5), 582-591. https://doi.org/10.1177/0956797620904450

Cheung, M. W. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology, 5*(1521). https://doi.org/https://doi.org/10.3389/fpsyg.2014.01521

Connolly, E. J., Schwartz, J. A., Nedelec, J. L., Beaver, K. M., & Barnes, J. C. (2015). Different slopes for different folks: Genetic influences on growth in delinquent peer association and delinquency during adolescence. *Journal of Youth and Adolescence, 44*(7), 1413-1427. https://doi.org/10.1007/s10964-015-0299-8

Cooper, R. M., & Zubek, J. P. (1958). Effects of enriched and restricted early environments on the learning ability of bright and dull rats. *Canadian Journal of Psychology, 12*(3), 159-164. https://doi.org/10.1037/h0083747

Darling, J. C., Bamidis, P. D., Burberry, J., & Rudolf, M. C. J. (2020). The first thousand days: Early, integrated and evidence-based approaches to improving child health: Coming to a population near you? *Archives of Disease in Childhood, 105*(9), 837–841. https://doi.org/http://dx.doi.org/10.1136/archdischild-2019-316929

Davis, O. S., Haworth, C. M., & Plomin, R. (2009). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science, 20*(10), 1301-1308. https://doi.org/10.1111/j.1467-9280.2009.02433.x

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Derks, E. M., Dolan, C. V., & Boomsma, D. I. (2006). A test of the equal environment assumption (EEA) in multivariate twin studies. *Twin Research and Human Genetics, 9*(3), 403-411. https://doi.org/10.1375/twin.9.3.403

Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of Clinical Psychology, 7*(1), 383-409. https://doi.org/10.1146/annurev-clinpsy-032210-104518

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346. https://doi.org/10.1037/0033-295X.108.2.346

Eaves, L. J., & Rao, D. C. (1984). The resolution of genotype × environment interaction in segregation analysis of nuclear families. *Genetic Epidemiology, 1*(3), 215-228. https://doi.org/10.1002/gepi.1370010302

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/https://doi.org/10.1136/bmj.315.7109.629

Elam, K. K., Harold, G. T., Neiderhiser, J. M., Reiss, D., Shaw, D. S., Natsuaki, M. N., Gaysina, D., Barrett, D., & Leve, L. D. (2014). Adoptive parent hostility and children's peer behavior problems: Examining the role of genetically informed child attributes on adoptive parent behavior. *Developmental Psychology, 50*(5), 1543-1552. https://doi.org/10.1037/a0035470

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

Fearon, P., Reiss, D., Leve, L. D., Shaw, D. S., Scaramella, L. V., Ganiban, J. M., & Neiderhiser, J. M. (2015). Child-evoked maternal negativity from 9 to 27 months: Evidence of gene–environment correlation and its moderation by marital distress. *Development and Psychopathology, 27*(4pt1), 1251-1265. https://doi.org/10.1017/S0954579414000868

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh, 52*(2), 399-433.

Flom, M., Wang, M., Uccello, K. J., & Saudino, K. J. (2018). Parent- and observer-rated positive affect in early childhood: Genetic overlap and environmental specificity. *Behavior Genetics, 48*(6), 432-439. https://doi.org/10.1007/s10519-018-9924-0

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*(2), 171. https://doi.org/10.1037/0033-2909.101.2.171

Ge, X., Conger, R. D., Cadoret, R. J., Neiderhiser, J. M., Yates, W., Troughton, E., & Stewart, M. A. (1996). The developmental interface between nature and nurture: A mutual influence model of child antisocial behavior and parent behaviors. *Developmental Psychology, 32*(4), 574-589. https://doi.org/10.1037/0012-1649.32.4.574

Harold, G. T., Leve, L. D., Barrett, D., Elam, K. K., Neiderhiser, J. M., Natsuaki, M. N., Shaw, D. S., Reiss, D., & Thapar, A. (2013). Biological and rearing mother influences on child ADHD symptoms: Revisiting the developmental interface between nature and nurture. *Journal of Child Psychology and Psychiatry, 54*(10), 1038-1046. https://doi.org/10.1111/jcpp.12100

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S. P., Kovas, Y., Corley, R. P., DeFries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*(11), 1112-1120. https://doi.org/10.1038/mp.2009.55

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry, 46*(3), 289. https://doi.org/https://doi.org/10.1111/j.1465-7295.2008.00163.x

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560. https://doi.org/https://doi.org/10.1136/bmj.327.7414.557

Horn, J. M. (1983). The Texas Adoption Project: Adopted children and their intellectual resemblance to biological and adoptive parents. *Child Development, 54*(2), 268-275. https://doi.org/10.2307/1129690

Hummer, R. A., & Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States. *Population bulletin, 68*(1), 1-16.

Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine, 37*(5), 615-626. https://doi.org/10.1017/S0033291706009524

Klahr, A. M., Burt, S. A., Leve, L. D., Shaw, D. S., Ganiban, J. M., Reiss, D., & Neiderhiser, J. M. (2017). Birth and adoptive parent antisocial behavior and parenting: A study of evocative gene–environment correlation. *Child Development, 88*(2), 505-513. https://doi.org/10.1111/cdev.12619

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*.

Koellinger, P. D., & Harden, K. P. (2018). Using nature to understand nurture. *Science, 359*(6374), 386-387. https://doi.org/10.1126/science.aar6429

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

Kosik, R., Mandell, G., Fan, A. P., Nguyen, T., Chen, J., & Eaton, W. (2018). The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study. *The European Journal of Psychiatry, 32*(2), 53-62. https://doi.org/10.1016/j.ejpsy.2018.01.001

Krapohl, E., Hannigan, L. J., Pingault, J.-B., Patel, H., Kadeva, N., Curtis, C., Breen, G., Newhouse, S. J., Eley, T. C., O'Reilly, P. F., & Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences, 114*(44), 11727-11732. https://doi.org/10.1073/pnas.1707178114

Kumsta, R., Kreppner, J., Kennedy, M., Knights, N., Rutter, M., & Sonuga-Barke, E. (2015). Psychological consequences of early global deprivation: An overview of findings from the English & Romanian Adoptees study. *European Psychologist, 20*(2), 138. https://doi.org/10.1027/1016-9040/a000227

Lejarraga, H. (2012). Growth in infancy and childhood: A pediatric approach. In N. Cameron & B. Bogin (Eds.), *Human Growth and Development* (Second ed., pp. 23–56). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-383882-7.00002-7

Lettre, G. (2011). Recent progress in the study of the genetics of height. *Human Genetics, 129*(5), 465-472. https://doi.org/10.1007/s00439-011-0969-x

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 22*(6), 716-727. https://doi.org/10.1017/thg.2019.66

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J. M., Natsuaki, M. N., & Reiss, D. (2013). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*(1), 412-423. https://doi.org/10.1017/thg.2012.126

MacGillivray, I., Campbell, D. M., & Thompson, B. (1988). *Twinning and twins*. Wiley.

Marceau, K., Horwitz, B. N., Narusyte, J., Ganiban, J. M., Spotts, E. L., Reiss, D., & Neiderhiser, J. M. (2013). Gene–environment correlation underlying the association between parental negativity and adolescent externalizing problems. *Child Development, 84*(6), 2031-2046. https://doi.org/10.1111/cdev.12094

McGue, M., Bouchard Jr., T. J., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (Vol. 1, pp. 59-76). American Psychological Association.

McGue, M., Keyes, M., Sharma, A., Elkins, I., Legrand, L., Johnson, W., & Iacono, W. G. (2007). The environments of adopted and non-adopted youth: Evidence on range restriction from the Sibling Interaction and Behavior Study (SIBS). *Behavior Genetics, 37*(3), 449-462. https://doi.org/10.1007/s10519-007-9142-7

Molenaar, D., & Dolan, C. V. (2014). Testing systematic genotype by environment interactions using item level data. *Behavior Genetics, 44*(3), 212-231. https://doi.org/10.1007/s10519-014-9647-9

Morris, T. T., Davies, N. M., Hemani, G., & Smith, G. D. (2019). Why are education, socioeconomic position and intelligence genetically correlated? *bioRxiv*, 630426. https://doi.org/10.1101/630426

NCD Risk Factor Collaboration. (2016). A century of trends in adult human height. *Elife, 5*. https://doi.org/10.7554/eLife.13410

Neale, M. C., & Stevenson, J. (1989). Rater bias in the EASI temperament scales: A twin study. *Journal of Personality and Social Psychology, 56*(3), 446-455. https://doi.org/10.1037//0022-3514.56.3.446

Nelson, C. A., 3rd, Zeanah, C. H., & Fox, N. A. (2019). How early experience shapes human development: The case of psychosocial deprivation. *Neural Plasticity, 2019*. https://doi.org/https://doi.org/10.1155/2019/1676285

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlskog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., Porteous, D. J., Freese, J., Herd, P., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., Hinds, D. A., Huber, K. E., Kleinman, A., Litterman, N. K., McCreight, J. C., McIntyre, M. H., Mountain, J. L., Northover, C. A. M., Pitts, S. J., Sathirapongsasuti, J. F., Sazonova, O. V., Shelton, J. F., Shringarpure, S., Tung, J. Y., Vacic, V., Wilson, C. H., Fontana, M. A., Pers, T. H., Rietveld, C. A., Chen, G.-B., Emilsson, V., Meddens, S. F. W., Pickrell, J. K., Thom, K., Timshel, P., de Vlaming, R., Abdellaoui, A., Ahluwalia, T. S., Bacelis, J., Baumbach, C., Bjornsdottir, G., Brandsma, J. H., Concas, M. P., Derringer, J., Galesloot, T. E., Girotto, G., Gupta, R., Hall, L. M., Harris, S. E., Hofer, E., Horikoshi, M., Huffman, J. E., Kaasik, K., Kalafati, I. P., Karlsson, R., Lahti, J., van der Lee, S. J., de Leeuw, C., Lind, P. A., Lindgren, K.-O., Liu, T., Mangino, M., Marten, J., Mihailov, E., Miller, M. B., van der Most, P. J., Oldmeadow, C., Payton, A., Pervjakova, N., Peyrot, W. J., Qian, Y., Raitakari, O., Rueedi, R., Salvi, E., Schmidt, B., Schraut, K. E., Shi, J., Smith, A. V., Poot, R. A., Pourcain, B. S., Teumer, A., Thorleifsson, G., Verweij, N., Vuckovic, D., Wellmann, J., Westra, H.-J., Yang, J., Zhao, W., Zhu, Z., Alizadeh, B. Z., Amin, N., Bakshi, A., Baumeister, S. E., Biino, G., Bønnelykke, K., Boyle, P. A., Campbell, H., Cappuccio, F. P., Davies, G., De Neve, J.-E., Deloukas, P., Demuth, I., Ding, J., Eibich, P., Eisele, L., Eklund, N., Evans, D. M., Faul, J. D., Feitosa, M. F., Forstner, A. J., Gandin, I., Gunnarsson, B., Halldórsson, B. V., Harris, T. B., Heath, A. C., Hocking, L. J., Holliday, E. G., Homuth, G., Horan, M. A., Hottenga, J.-J., de Jager, P. L., Joshi, P. K., Jugessur, A., Kaakinen, M. A., Kähönen, M., Kanoni, S., Keltigangas-Järvinen, L., Kiemeney, L. A. L. M., Kolcic, I., Koskinen, S., Kraja, A. T., Kroh, M., Kutalik, Z., Latvala, A., Launer, L. J., Lebreton, M. P., Levinson, D. F., Lichtenstein, P., Lichtner, P., Liewald, D. C. M., Loukola, A., Madden, P. A., Mägi, R., Mäki-Opas, T., Marioni, R. E., Marques-Vidal, P., Meddens, G. A., McMahon, G., Meisinger, C., Meitinger, T., Milaneschi, Y., Milani, L., Montgomery, G. W., Myhre, R., Nelson, C. P., Nyholt, D. R., Ollier, W. E. R., Palotie, A., Paternoster, L., Pedersen, N. L., Petrovic, K. E., Räikkönen, K., Ring, S. M., Robino, A., Rostapshova, O., Rudan, I., Rustichini, A., Salomaa, V., Sanders, A. R., Sarin, A.-P., Schmidt, H., Scott, R. J., Smith, B. H., Smith, J. A., Staessen, J. A., Steinhagen-Thiessen, E., Strauch, K., Terracciano, A., Tobin, M. D., Ulivi, S., Vaccargiu, S., Quaye, L., van Rooij, F. J. A., Venturini, C., Vinkhuyzen, A. A. E., Völker, U., Völzke, H., Vonk, J. M., Vozzi, D., Waage, J., Ware, E. B., Willemsen, G., Attia, J. R., Bennett, D. A., Berger, K., Bertram, L., Bisgaard, H., Boomsma, D. I., Borecki, I. B., Bültmann, U., Chabris, C. F., Cucca, F., Cusi, D., Deary, I. J., Dedoussis, G. V., van Duijn, C. M., Eriksson, J. G., Franke, B., Franke, L., Gasparini, P., Gejman, P. V., Gieger, C., Grabe, H.-J., Gratten, J., Groenen, P. J. F., Gudnason, V., van der Harst, P., Hoffmann, W., Hyppönen, E., Iacono, W. G., Jacobsson, B., Järvelin, M.-R., Jöckel, K.-H., Kaprio, J., Kardia, S. L. R., Lehtimäki, T., Lehrer, S. F., Martin, N. G., McGue, M., Metspalu, A., Pendleton, N., Penninx, B. W. J. H., Perola, M., Pirastu, N., Pirastu, M., Polasek, O., Posthuma, D., Power, C., Province, M. A., Samani, N. J., Schlessinger, D., Schmidt, R., Sørensen, T. I. A., Spector, T. D., Stefansson, K., Thorsteinsdottir, U., Thurik, A. R., Timpson, N. J., Tiemeier, H., Uitterlinden, A. G., Vitart, V., Vollenweider, P., Weir, D. R., Wilson, J. F., Wright, A. F., Conley, D. C., Krueger, R. F., Smith, G. D., Hofman, A., Laibson, D. I., Medland, S. E., Yang, J., Esko, T., Watson, C., Jala, J., Conley, D., Koellinger, P. D., Johannesson, M., Laibson, D., Meyer, M. N., Lee, J. J., Kong, A., Yengo, L., Cesarini, D., Turley, P., Visscher, P. M., Beauchamp, J. P., andMe Research, T., Social Science Genetic Association, C., & LifeLines Cohort, S. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics, 54*(4), 437-449. https://doi.org/10.1038/s41588-022-01016-z

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal, 372*. https://doi.org/https://doi.org/10.1136/bmj.n71

Papageorgiou, K. A., & Ronald, A. (2017). The genetic basis of psychological traits in infancy: Implications for understanding the causes of developmental psychopathology. In L. C. Centifanti & D. M. Williams (Eds.), *The Wiley handbook of developmental psychopathology* (pp. 235–258). Wiley Blackwell. https://doi.org/https://doi.org/10.1002/9781118554470.ch11

Pappa, I., Szekely, E., Mileva-Seitz, V. R., Luijk, M. P., Bakermans-Kranenburg, M. J., van, I. M. H., & Tiemeier, H. (2015). Beyond the usual suspects: A multidimensional genetic exploration of infant attachment disorganization and security. *Attachment and Human Development, 17*(3), 288–301. https://doi.org/https://doi.org/10.1080/14616734.2015.1037316

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A.

C., Kalungi, A., Koen, N., Majara, L., Schwarz, E., Smoller, J. W., Stahl, E. A., Sullivan, P. F., Vassos, E., Mowry, B., Prieto, M. L., Cuellar-Barboza, A., Bigdeli, T. B., Edenberg, H. J., Huang, H., & Duncan, L. E. (2019). Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. *Cell, 179*(3), 589-603. https://doi.org/10.1016/j.cell.2019.08.051

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science, 10*(3), 282-306. https://doi.org/10.1177/1745691615577701

Pike, A., McGuire, S., Hetherington, E., Reiss, D., & Plomin, R. (1996). Family environment and adolescent depressive symptoms and antisocial behavior: A multivariate genetic analysis. *Developmental Psychology, 32*(4), 590-603.

Plomin, R. (2011). Commentary: Why are children in the same family so different? Non-shared environment three decades later. *International Journal of Epidemiology, 40*(3), 582-592. https://doi.org/10.1093/ije/dyq144

Plomin, R. (2018). *Blueprint: How DNA makes us who we are*. Allen Lane.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on “environmental” measures. *Behavioral and Brain Sciences, 14*(3), 373-386. https://doi.org/10.1017/S0140525X00070278

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10*(1), 1-16. https://doi.org/10.1017/S0140525X00055941

Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry, 20*(1), 98-108. https://doi.org/10.1038/mp.2014.105

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science, 11*(1), 3-23. https://doi.org/10.1177/1745691615617439

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*(2), 309-322.

Plomin, R., Fulker, D. W., Corley, R. P., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 Years: A parent-offspring adoption study. *Psychological Science, 8*(6), 442-447. https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

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

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics, 47*(7), 702-709. https://doi.org/10.1038/ng.3285

R Core Team. (2020). *A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rhea, S.-A., Bricker, J. B., Wadsworth, S. J., & Corley, R. P. (2013). The Colorado Adoption Project. *Twin Research and Human Genetics, 16*(1), 358-365. https://doi.org/10.1017/thg.2012.109

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., Lewis, G., Harold, G. T., & Thapar, A. (2013). Examining the role of passive gene–environment correlation in childhood depression using a novel genetically sensitive design. *Development and Psychopathology, 25*(1), 37-50. https://doi.org/10.1017/S0954579412000880

Robinson, M. R., Kleinman, A., Graff, M., Vinkhuyzen, A. A. E., Couper, D., Miller, M. B., Peyrot, W. J., Abdellaoui, A., Zietsch, B. P., Nolte, I. M., van Vliet-Ostaptchouk, J. V., Snieder, H., Alizadeh, B. Z., Boezen, H. M., Franke, L., van der Harst, P., Navis, G., Rots, M., Snieder, H., Swertz, M., Wolffenbuttel, B. H. R., Wijmenga, C., Abecasis, G. R., Absher, D., Alavere, H., Albrecht, E., Allen, H. L., Almgren, P., Amin, N., Amouyel, P., Anderson, D., Arnold, A. M., Arveiler, D., Aspelund, T., Asselbergs, F. W., Assimes, T. L., Atalay, M., Attwood, A. P., Atwood, L. D., Bakker, S. J. L., Balkau, B., Balmforth, A. J., Barlassina, C., Barroso, I., Basart, H., Bauer, S., Beckmann, J. S., Beilby, J. P., Bennett, A. J., Ben-Shlomo, Y., Bergman, R. N., Bergmann, S., Berndt, S. I., Biffar, R., Di Blasio, A. M., Boehm, B. O., Boehnke, M., Boeing, H., Boerwinkle, E., Bolton, J. L., Bonnefond, A., Bonnycastle, L. L., Boomsma, D. I., Borecki, I. B., Bornstein, S. R., Bouatia-Naji, N., Boucher, G., Bragg-Gresham, J. L., Brambilla, P., Bruinenberg, M., Buchanan, T. A., Buechler, C., Cadby, G., Campbell, H., Caulfield, M. J., Cavalcanti-Proença, C., Cesana, G., Chanock, S. J., Chasman, D. I., Chen, Y. I., Chines, P. S., Clegg, D. J., Coin, L., Collins, F. S., Connell, J. M., Cookson, W., Cooper, M. N., Croteau-Chonka, D. C., Cupples, L. A., Cusi, D., Day, F. R., Day, I. N. M., Dedoussis, G. V., Dei, M., Deloukas, P., Dermitzakis, E. T., Dimas, A. S., Dimitriou, M., Dixon, A. L., Dörr, M., van Duijn, C. M., Ebrahim, S., Edkins, S., Eiriksdottir, G., Eisinger, K., Eklund, N., Elliott, P., Erbel, R., Erdmann, J., Erdos, M. R., Eriksson, J. G., Esko, T., Estrada, K., Evans, D. M., de Faire, U., Fall, T., Farrall, M., Feitosa, M. F., Ferrario, M. M., Ferreira, T., Ferrières, J., Fischer, K., Fisher, E., Fowkes, G., Fox, C. S., Franke, L., Franks, P. W., Fraser, R. M., Frau, F., Frayling, T., Freimer, N. B., Froguel, P., Fu, M., Gaget, S., Ganna, A., Gejman, P. V., Gentilini, D., Geus, E. J. C., Gieger, C., Gigante, B., Gjesing, A. P., Glazer, N. L., Goddard, M. E., Goel, A., Grallert, H., Gräßler, J., Grönberg, H., Groop, L. C., Groves, C. J., Gudnason, V., Guiducci, C., Gustafsson, S., Gyllensten, U., Hall, A. S., Hall, P., Hallmans, G., Hamsten, A., Hansen, T., Haritunians, T., Harris, T. B., van der Harst, P., Hartikainen, A., Hassanali, N., Hattersley, A. T., Havulinna, A. S., Hayward, C., Heard-Costa, N. L., Heath, A. C., Hebebrand, J., Heid, I. M., den Heijer, M., Hengstenberg, C., Herzig, K., Hicks, A. A., Hingorani, A., Hinney, A., Hirschhorn, J. N., Hofman, A., Holmes, C. C., Homuth, G., Hottenga, J., Hovingh, K. G., Hu, F. B., Hu, Y., Huffman, J. E., Hui, J., Huikuri, H., Humphries, S. E., Hung, J., Hunt, S. E., Hunter, D., Hveem, K., Hyppönen, E., Igl, W., Illig, T., Ingelsson, E., Iribarren, C., Isomaa, B., Jackson, A. U., Jacobs, K. B., James, A. L., Jansson, J., Jarick, I., Jarvelin, M., Jöckel, K. H., Johansson, Å., Johnson, T., Jolley, J., Jørgensen, T., Jousilahti, P., Jula, A., Justice, A. E., Kaakinen, M., Kähönen, M., Kajantie, E., Kanoni, S., Kao, W. H. L., Kaplan, L. M., Kaplan, R. C., Kaprio, J., Kapur, K., Karpe, F., Kathiresan, S., Kee, F., Keinanen-Kiukaanniemi, S. M., Ketkar, S., Kettunen, J., Khaw, K., Kiemeney, L. A., Kilpeläinen, T. O., Kinnunen, L., Kivimaki, M., Kivmaki, M., Van der Klauw, M. M., Kleber, M. E., Knowles, J. W., Koenig, W., Kolcic, I., Kolovou, G., König, I. R., Koskinen, S., Kovacs, P., Kraft, P., Kraja, A. T., Kristiansson, K., Krjutškov, K., Kroemer, H. K., Krohn, J. P., Krzelj, V., Kuh, D., Kulzer, J. R., Kumari, M., Kutalik, Z., Kuulasmaa, K., Kuusisto, J., Kvaloy, K., Laakso, M., Laitinen, J. H., Lakka, T. A., Lamina, C., Langenberg, C., Lantieri, O., Lathrop, G. M., Launer, L. J., Lawlor, D. A., Lawrence, R. W., Leach, I. M., Lecoeur, C., Lee, S. H., Lehtimäki, T., Leitzmann, M. F., Lettre, G., Levinson, D. F., Li, G., Li, S., Liang, L., Lin, D., Lind, L., Lindgren, C. M., Lindström, J., Liu, J., Liuzzi, A., Locke, A. E., Lokki, M., Loley, C., Loos, R. J. F., Lorentzon, M., Luan, J., Luben, R. N., Ludwig, B., Madden, P. A., Mägi, R., Magnusson, P. K. E., Mangino, M., Manunta, P., Marek, D., Marre, M., Martin, N. G., The LifeLines

Cohort Study, & Genetic Investigation of Anthropometric Traits Consortium. (2017). Genetic evidence of assortative mating in humans. *Nature Human Behaviour, 1*(1), 0016. https://doi.org/10.1038/s41562-016-0016

Ronald, A., Happe, F., & Plomin, R. (2008). A twin study investigating the genetic and environmental aetiologies of parent, teacher and child ratings of autistic-like traits and their overlap. *European Child and Adolescent Psychiatry, 17*(8), 473–483. https://doi.org/https://doi.org/10.1007/s00787-008-0689-5

Ronalds, G. A., De Stavola, B. L., & Leon, D. A. (2005). The cognitive cost of being a twin: Evidence from comparisons within families in the Aberdeen children of the 1950s cohort study. *British Medical Journal, 331*(7528), 1306. https://doi.org/https://doi.org/10.1136/bmj.38633.594387.3A

Rutter, M. (1998, May). Developmental catch-up, and deficit, following adoption after severe global early privation. *The Journal of Child Psychology and Psychiatry, 39*(4), 465–476. https://www.ncbi.nlm.nih.gov/pubmed/9599775

Rutter, M. (2000). Psychosocial influences: Critiques, findings, and research needs. *Development and Psychopathology, 12*(3), 375-405. https://doi.org/10.1017/S0954579400003072

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C. M., & Reichenberg, A. (2014). The familial risk of autism. *JAMA, 311*(17), 1770-1777. https://doi.org/10.1001/jama.2014.4144

Saudino, K. J., Ronald, A., & Plomin, R. (2005). The etiology of behavior problems in 7-year-old twins: Substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. *Journal of Abnormal Child Psychology, 33*(1), 113–130. https://doi.org/https://doi.org/10.1007/s10802-005-0939-7

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424-435. https://doi.org/10.2307/1129703

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., Tomizawa, R., Watanabe, M., Sakai, N., Rebato, E., Busjahn, A., Tyler, J., Hopper, J. L., Ordoñana, J. R., Sánchez-Romera, J. F., Colodro-Conde, L., Calais-Ferreira, L., Oliveira, V. C., Ferreira, P. H., Medda, E., Nisticò, L., Toccaceli, V., Derom, C. A., Vlietinck, R. F., Loos, R. J. F., Siribaddana, S. H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Duncan, G. E., Buchwald, D., Tynelius, P., Rasmussen, F., Tan, Q., Zhang, D., Pang, Z., Magnusson, P. K. E., Pedersen, N. L., Dahl Aslan, A. K., Hwang, A. E., Mack, T. M., Krueger, R. F., McGue, M., Pahlen, S., Brandt, I., Nilsen, T. S., Harris, J. R., Martin, N. G., Medland, S. E., Montgomery, G. W., Willemsen, G., Bartels, M., van Beijsterveldt, C. E. M., Franz, C. E., Kremen, W. S., Lyons, M. J., Silberg, J. L., Maes, H. H., Kandler, C., Nelson, T. L., Whitfield, K. E., Corley, R. P., Huibregtse, B. M., Gatz, M., Butler, D. A., Tarnoki, A. D., Tarnoki, D. L., Park, H. A., Lee, J., Lee, S. J., Sung, J., Yokoyama, Y., Sørensen, T. I. A., Boomsma, D. I., & Kaprio, J. (2020). Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. *Scientific Reports, 10*(1), 12681. https://doi.org/10.1038/s41598-020-69526-6

Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2019). Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry, 24*(3), 409-420. https://doi.org/10.1038/s41380-017-0010-4

Stoolmiller, M. (1999). Implications of the restricted range of family environments for estimates of heritability and nonshared environment in behavior-genetic adoption studies. *Psychological Bulletin, 125*(4), 392-409. https://doi.org/10.1037/0033-2909.125.4.392

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. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA, 283*(15), 2008-2012. https://doi.org/10.1001/jama.283.15.2008

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics, 20*(8), 467-484. https://doi.org/10.1038/s41576-019-0127-1

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

Tucker-Drob, E. M., & Bates, T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138-149. https://doi.org/10.1177/0956797615612727

Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene–environment transactions. *Developmental Science, 15*(2), 250-259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 9*(5), 160-164. https://doi.org/10.1111/1467-8721.00084

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, II. (2003). Socioeconomic status modifies heritability of IQ in young children.

Psychological Science, 14(6), 623-628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

Verhulst, B. (2017). A power calculator for the classical twin design. *Behavior Genetics, 47*(2), 255-261. https://doi.org/10.1007/s10519-016-9828-9

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

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

Wang, B., Baldwin, J. R., Schoeler, T., Cheesman, R., Barkhuizen, W., Dudbridge, F., Bann, D., Morris, T. T., & Pingault, J. B. (2021). Robust genetic nurture effects on education: A systematic review and meta-analysis based on 38,654 families across 8 cohorts. *American Journal of Human Genetics, 108*(9), 1780-1791. https://doi.org/10.1016/j.ajhg.2021.07.010

Wang, Y., Tsuo, K., Kanai, M., Neale, B. M., & Martin, A. R. (2022). Challenges and opportunities for developing more generalizable polygenic risk scores. *Annual Review of Biomedical Data Science, Advance online publication*. https://doi.org/10.1146/annurev-biodatasci-111721-074830

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd, L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745-1761. https://doi.org/10.1111/cdev.13329

World Health Organization. (2007). *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY*. World Health Organization.

Wright, S. (1921). Systems of mating. *Genetics, 6*(2), 111-178. https://doi.org/10.1093/genetics/6.2.167

Chapter 2: Heritability of Psychological Traits and Developmental Milestones in Infancy: A Systematic Review and Meta-Analysis of Twin Studies

A version of this study was published in *JAMA Network Open* (Austerberry, Mateen, et al., 2022).

2.1 Introduction

Infancy represents the most rapid period of postnatal growth and development (Lejarraga, 2012), and research suggests that it is a sensitive period for a wide range of psychological and developmental milestones (Kumsta et al., 2015; Nelson et al., 2019; Rutter, 1998). Investment in early childhood is argued to be one of the most effective economic strategies through promoting long-term socioeconomic and health outcomes (Heckman, 2008). Investment before age two years, in particular, appears to yield the greatest rate of return for investment (Heckman, 2008). This is reflected in an increasing policy focus globally on the first thousand and one days from conception to age two years (Darling et al., 2020).

Variation in complex traits appears to be influenced by a combination of genetic and environmental differences (Polderman et al., 2015). Consequently, to gain understanding of the development of traits in infancy, it is important to draw on literature examining genetic and environmental sources of infant trait variation. The quantitative genetic method most widely and comprehensively performed in infancy is the classical twin design, which has been used for over a century partition phenotypic variance into additive genetic variance (heritability) and variance in the shared and nonshared environment. Family studies comparing biologically related siblings or parent-offspring are typically unable to separate genetics from shared environment. In contrast, the classical twin design can provide separate estimates of heritability (the proportion of trait variation explained by genetic differences) and shared and nonshared environment. Twin studies are more feasible than adoption studies (which compare degree of resemblance between adoptees and their birth parents with resemblance between adoptees and their adoptive parents) to conduct at scale during infancy because adoption often occurs later in childhood. This has resulted in a far smaller and less comprehensive body of evidence in infancy from adoption studies than twin studies. The molecular genetic literature on infant traits is also small; the first genome-wide association study of infant traits was only recently conducted (Pappa et al., 2015), and most molecular genetic studies in infancy have used candidate gene association methods, which in general have failed to yield replicable findings (Papageorgiou & Ronald, 2017).

A landmark meta-analysis (Polderman et al., 2015), synthesizing virtually all twin studies of complex traits (predominantly psychiatric, metabolic and cognitive) found a heritability of 49% across the lifespan when all traits and age groups were combined. The analysis combined data from infants and older children, calculating pooled estimates for children aged 0 to 11 years. Infancy is a rapid and sensitive period of development that deserves special focus. To address this, I conducted the first (to my knowledge) meta-analysis of twin studies of psychological and developmental functioning, disability, and health in infancy (birth to age 2 years), calculating pooled estimates of heritability and shared and nonshared environment.

2.2 Methods

This study protocol was registered with PROSPERO (record number: CRD42019151532), and the systematic review and meta-analysis were performed in line with the Preferred Reporting Items for Systematic Reviews and the Meta-analyses (PRISMA) 2020 statement and Meta-analysis Of Observational Studies in Epidemiology (MOOSE) proposal for reporting (Page, McKenzie, et al., 2021; Stroup et al., 2000). Given that the review involved the synthesis of anonymized information available in the public domain, it was exempt, according to the University College London Research Ethics Committee (UCL REC) regulations, from requirements for ethics review by the UCL REC and the need for informed consent.

2.2.1 Search Strategy

I searched PubMed and PsycINFO databases on November 30, 2018; February 5, 2020; and February 11, 2021, for twin studies (a genetically informed design described in the Supplementary Methods subsection of Appendix A) of psychological traits and developmental milestones in infancy, using the search terms in Table A1 (Appendix A). I imported the search results into Endnote (Version 9). I reviewed duplicates identified by Endnote, deleting true duplicates, and screened the titles and abstracts of all identified records against the inclusion and exclusion criteria in Table A2 (Appendix A). I retrieved the full texts for all nonexcluded records, and these, along with references of included publications, were screened by me and a DClinPsy student (M.M.) Uncertainty about whether publications met inclusion criteria was resolved with my PhD supervisors, P.F. and A.R.

2.2.2 Quality Assessment and Data Extraction

I rated included publications using an adaptation for twin studies of the Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields for Quantitative Studies (Kmet et al., 2004), detailed in the Supplementary Methods (Appendix A). M.M. and I extracted the information in Table A1 (Appendix A) from each included publication. If publications reported overlapping data, the estimate with the larger sample size (or, if sample sizes were identical, the most recently published estimate) was retained for meta-analysis (Supplementary Methods (Appendix A).

2.2.3 Classification of Phenotypes

I classified the phenotypes using the World Health Organization *International Classification of Functioning, Disability and Health for Children and Youth* (*ICF-CY*) (World Health Organization, 2007). Uncertainty about which *ICF-CY* category I should place a phenotype in was resolved through discussion with M.M., P.F. and A.R. Phenotypes were excluded from the meta-analysis if they could not be categorized or were in categories containing data from fewer than five independent samples.

2.2.4 Statistical Analysis

Before running the meta-analyses, I prepared the extracted data. If studies reported only heritability, shared and nonshared environment (ACE) estimates or both ACE estimates and concordances (not correlations), I used Falconer's formulas to convert the ACE estimates into correlations (Falconer, 1960). Variances were then calculated using the within-twin correlations (rMZ and rDZ) and sample sizes. The correlations and variances were then used in the meta-analyses. In the metafor package version 2.4-0 (Viechtbauer, 2010), using the R statistical software version 4.0.2 (R Core Team, 2020), I conducted two three-level multilevel random-effects models (incorporating sampling variance, within-cohort variance in outcome measurements, and between-cohort variance) on twin correlations weighted by sample size from the 10 categories of the *ICF-CY* containing data from five or more twin cohorts (Supplementary Methods,

Appendix A). Zygosity was included as a moderator, with the dizygotic (DZ) group coded as the reference category in the first model, to obtain a pooled monozygotic (MZ) twin correlation (r_{MZ}) and standard error. The second model was identical but re-parameterized with the MZ group as the reference category, producing a pooled DZ twin correlation (r_{DZ}) and standard error. To allow for differences in variability in MZ and DZ subsets, models had a random error structure creating separate study-level and outcome error terms for MZ and DZ twins.

Next, after running the multilevel random effects meta-analyses, I calculated ACE estimates for the 10 *ICF-CY* categories by running meta-analytic SEM models using the R package metaSEM version 1.2.5.1 (Cheung, 2015)—see Supplementary Methods (Appendix A). These models were estimated using the correlations and squared standard errors (variances) from each of the two multilevel meta-analyses (one to estimate the pooled within-twin correlation and standard error for MZ twins and one to estimate the pooled within-twin correlation and standard error for DZ twins), allowing parameters and their confidence intervals to be estimated taking account of between-study heterogeneity. Ninety-five percent confidence intervals for the ACE parameter estimates were obtained using likelihood-based confidence interval estimation in OpenMx.

I produced forest plots for the analyses using the R package metafor version 2.4-0 (Viechtbauer, 2010). I calculated I^2 for each of the three levels in the multilevel models. According to Cochrane guidelines, $I^2 \leq 40\%$ suggests low heterogeneity, 30–60% suggests moderate heterogeneity and ≥ 50% indicates substantial/considerable heterogeneity (Higgins et al., 2003). In an attempt to reduce heterogeneity, I repeated the analysis steps in 10 narrower *ICF-CY* subcategories (with data from ≥ 5 samples) and 3 *ICF-CY* categories (with separate data from parents and observers from ≥ 5 samples) by parent and observer subgroups (for 6 meta-analyses in total) given that differences in rater have been shown to influence heritability estimates (Ronald et al., 2008; Saudino et al., 2005).

I ran Egger's tests of publication bias, using the standard error as the predictor, and created funnel plots, plotting effect sizes against standard errors (Egger et al., 1997). Egger tests of publication bias were 2-sided and were considered significant at $P < .05$. In line with Cochrane recommendations, publication bias tests were only run on estimates in trait categories containing at least 10 estimates (Higgins et al., 2003). Data analysis was conducted March through September 2021.

2.3 Results

I identified 5047 publications (4675 in databases and 372 in references). After duplicate removal and screening, I included 139 publications (citations for which are in Appendix A), containing data on 79,044 twin pairs (31,053 MZ, 47,991 DZ), 52 twin cohorts, 21 countries, and 6 continents between 1972 and 2020. M.M. and I extracted 2279 estimates (twin correlations or ACE estimates, including 1097 estimates from MZ twins and 1182 from DZ) on 377 phenotypes, organized into 17 categories and 28 sub-categories of the *ICF-CY*. Data from 33 publications (citations for which are in Appendix A) included in the systematic review were excluded from the meta-analysis. Detailed information on search results, phenotype categorization and excluded data is provided in the Supplementary Results, Tables 1A and 4A–6A, and Figures 1A–3A (Appendix A).

2.3.1 Meta-Analysis Results

Analysis of Phenotypes by Category. Among 10 categories of infant psychological and developmental functioning, disability and health displayed in Figure 2.1 and defined in the *ICF-CY* (World Health Organization, 2007), contained data from enough independent samples for meta-analysis (≥ 5 samples). Results are reported in Table 2.1 and Figure 2.2. Forest plots for these meta-analyses are reported in Figures 4A–13A in Appendix A. More twin samples used in these meta-analyses contained parent-reported data ($k_{\text{cohort}} = 22$) than observer-rated data ($k_{\text{cohort}} = 12$)—see Table 7A and the Supplementary Results (Appendix A).

Heritability. The highest heritability estimate was for ‘psychomotor functions’ ($h^2_{\text{pooled}} = .59$; 95% CI [.25–.79] $p < .001$), followed by ‘attention functions’ ($h^2_{\text{pooled}} = .48$; 95% CI [.17–.71] $p = .002$), ‘complex interpersonal interactions’ ($h^2_{\text{pooled}} = .44$; 95% CI [.15–.75] $p = .003$), ‘family relationships’ ($h^2_{\text{pooled}} = .41$; 95% CI [.06–.71] $p = .022$) and ‘emotional functions’ ($h^2_{\text{pooled}} = .40$; 95% CI [.16–.64] $p = .001$). Remaining categories had modest estimates with 95% confidence intervals (CIs) above 0 (h^2_{pooled} range: .24–.38), apart from ‘mental functions of language’ and ‘sleep functions’, which had CIs overlapping 0 (h^2_{pooled} .24 and .35, respectively) (Table 2.1).

Shared Environment. ‘Mental functions of language’ (c^2_{pooled} = .59; 95% CI [.24–.86] p = .001), ‘growth maintenance functions’ (c^2_{pooled} = .46; 95% CI [.37–.54] p < .001), ‘basic cognitive functions’ (c^2_{pooled} = .45; 95% CI [.21–.69] p < .001) and ‘sleep functions’ (c^2_{pooled} = .45; 95% CI [.16–.74] p = .002) had high shared environment estimates. ‘Complex interpersonal interactions’ had a modest estimate (c^2_{pooled} = .27; 95% CI [.04–.51] p = .021), and estimates for ‘psychomotor functions’, ‘attention functions’, ‘emotional functions’, ‘family relationships’, and ‘basic interpersonal interactions’ had CIs overlapping 0 (c^2_{pooled} range: .07–.21) (Table 2.1).

Nonshared Environment. Categories with the highest nonshared environment estimates were ‘emotional functions’ (e^2_{pooled} = .42; 95% CI [.33–.50] p < .001), ‘family relationships’ (e^2_{pooled} = .42; 95% CI [.30–.55] p < .001), ‘basic interpersonal interactions’ (e^2_{pooled} = .41; 95% CI [.30–.52] p < .001), and ‘attention functions’ (e^2_{pooled} = .40; 95% CI [.29–.51] p < .001). The remaining categories all had modest estimates with CIs above 0 (e^2_{pooled} range: .18–.33) (Table 2.1).

Heterogeneity. Sampling variance contributed little to the total variance of each phenotypic category ($I^2_{\text{Level 1}}$ range: 0.19%–12.44%, Table 2.1). Within-cohort heterogeneity (i.e., differences across measures within a domain and within a cohort) contributed substantially to the total variance in ‘mental functions of language’, ‘emotional functions’, and ‘growth functions’ ($I^2_{\text{Level 2}}$ range: 58.59%–69.93%) and between-cohort heterogeneity contributed a low or moderate amount ($I^2_{\text{Level 3}}$ range: 24.73%–38.74%) to these outcomes. The remaining 7 categories each had substantial between-cohort heterogeneity ($I^2_{\text{Level 3}}$ range: 56.56%–75.28%) and low or moderate within-cohort heterogeneity ($I^2_{\text{Level 2}}$ range: 23.59%–42.04%) (Table 2.1).

Analysis of Phenotypes by Sub-Category and Rater. To reduce heterogeneity, I analysed 10 subcategories of the *ICF-CY* (with data from ≥ 5 samples) and 3 phenotypic categories (with separate parent and observer data from ≥ 5 samples) by rater (for 6 subgroups: 3 parent-report and 3 observer-report). Full findings are reported in Appendix A (Tables 8A and 9A and the Results subsection). Parent rated phenotypes in the 3 examined categories

(‘Psychomotor Functions’ and ‘Emotional Functions’ and ‘Basic Interpersonal Interactions’) had higher heritability and lower nonshared estimates than observer ratings and comparable shared environment estimates.

2.3.2 Publication Bias

Possible publication bias was detected in the unexpected direction across all categories. Findings are in the Supplementary Results subsection, Tables 10A and 11A, and Figures 14A–18A (Appendix A).

2.3.3 Quality Assessment

Quality assessment results are displayed in Figure 19A (Appendix A). The mean score for the 106 publications included in the meta-analysis (the citations for which are in Appendix A) was 75.58% (SD, 13.83%).

Table 2.1 Multilevel Random Effects Models of Phenotypic Categories

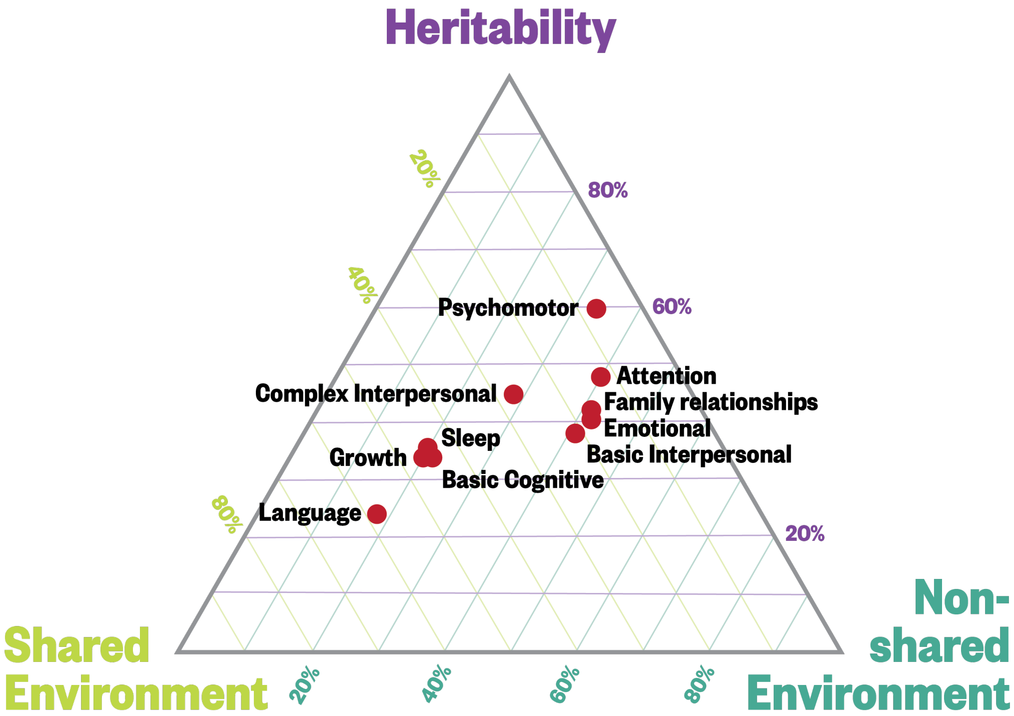
ICF-CY category	k_{cohort}	$k_{estimate}$	n_{MZ}	n_{DZ}	r_{MZ}	r_{MZ}	r_{DZ}	r_{DZ}	h^2	h^2	h^2 p	c^2	c^2	c^2 p	e^2	e^2	e^2 p	I^2	I^2	I^2
					pooled	*95% CI*	*pooled*	*95% CI*	*pooled*	*95% CI*		*pooled*	*95% CI*		*pooled*	*95% CI*		Level 1	Level 2	Level 3
b134 Sleep	7	49	1923	4044	.80	.67–.93	.63	.49–.76	.35	0–.73	.058	.45	.16–.74	.002	.20	.07–.33	.003	0.29	25.69	74.02
b140 Attention	10	175	3011	6137	.60	.49–.71	.36	.25–.47	.48	.17–.71	.002	.12	0–.37	.333	.40	.29–.51	<.001	12.44	30.26	57.29
b147 Psychomotor	13	151	3109	6105	.67	.55–.79	.37	.25–.49	.59	.25–.79	.001	.07	0–.35	.599	.33	.22–.45	<.001	1.03	23.69	75.28
b152 Emotional	14	216	1756	3633	.58	.50–.66	.38	.30–.46	.40	.16–.64	.001	.18	0–.38	.062	.42	.34–.50	<.001	6.48	68.79	24.73
b163 Basic cognitive	5	47	2636	5371	.79	.68–.89	.62	.51–.73	.34	.04–.64	.026	.45	.21–.69	<.001	.21	.11–.32	<.001	1.41	42.04	56.56
b167 Language	5	96	2232	2853	.82	.67–.98	.71	.55–.86	.24	0–.68	.284	.59	.24–.86	.001	.18	.02–.33	.023	0.19	69.93	29.88
b560 Growth	24	465	16653	21874	.80	.76–.83	.63	.59–.67	.34	.23–.45	<.001	.46	.37–.54	<.001	.20	.17–.24	<.001	2.67	58.59	38.74
d710 Basic interpersonal	18	356	4207	8037	.59	.48–.70	.40	.29–.51	.38	.05–.70	.021	.21	0–.48	.102	.41	.30–.52	<.001	1.62	23.59	74.79
d720 Complex interpersonal	10	73	3244	5117	.72	.61–.82	.49	.39–.60	.44	.15–.75	.003	.27	.04–.51	.021	.29	.18–.39	<.001	1.98	40.08	57.93
d760 Family relationships	7	29	678	1546	.58	.45–.71	.37	.24–.50	.41	.06–.71	.022	.17	0–.45	.237	.42	.30–.55	<.001	3.12	39.60	57.28

Note. k_{cohort} = number of independent twin cohorts. $k_{estimate}$ = number of estimates (twin correlations). n_{MZ} = number of monozygotic (MZ) twin pairs. n_{DZ} = number of dizygotic (DZ) twin pairs. r_{MZ} = MZ twin correlation. *95% CI* = 95% confidence interval. r_{DZ} = DZ twin correlation. h^2 = heritability. p = p-value. c^2 = shared environment. e^2 = nonshared environment. I^2 = heterogeneity. Level 1 = sampling variance. Level 2 = within-cohort variance in outcome measurement. Level 3 = between-cohort variance.

Figure 2.1 International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY) Categories Included in the Main Meta-Analysis

ICF-CY category and definition	Example phenotypes
b134 Sleep functions “General mental functions of periodic, reversible, and selective physical and mental disengagement from one’s immediate environment accompanied by characteristic physiological changes.” (WHO, 2007, p. 51)	Nocturnal sleep duration Night awakenings Sleep problems
b140 Attention functions “Specific mental functions of focusing on an external stimulus or internal experience for the required period of time” (WHO, 2007, p. 51)	Attention problems Task orientation Spectral amplitude during visual attention
b147 Psychomotor functions “Specific mental functions of control over both motor and psychological events at the body level.” (WHO, 2007, p.53)	Activity level Fine motor Sitting without support
b152 Emotional functions “Specific mental functions related to the feeling and affective components of the mind.” (WHO, 2007, p.54)	Resistance to soothing Fearfulness Positive attect
b163 Basic cognitive functions “Mental functions involved in acquisition of knowledge about objects, events and experiences; and the organization and application of that knowledge in tasks requiring mental activity.” (WHO, 2007, p.54)	General cognitive ability Nonverbal cognitive development Primary cognition
b167 Mental functions of language “Specific mental functions of recognizing and using signs, symbols and other components of a language.” (WHO, 2007, p.58)	Reception of language Expressive vocabulary Late language acquisition
b560 Growth maintenance functions “Functions of attaining expected growth milestones according to contextually adjusted normative auxological parameters.” (WHO, 2007, p.88)	BMI Head circumference Weight gain
d710 Basic interpersonal interactions “Interacting with people in a contextually and socially appropriate manner, such as by showing consideration and esteem when appropriate, or responding to the feelings of others.” (WHO, 2007, p.173)	Disregard for others Reciprocal social behaviour Shyness
d720 Complex interpersonal interactions “Maintaining and managing interactions with other people, in a contextually and socially appropriate manner, such as by regulating emotions and impulses, controlling verbal and physical aggression, acting independently in social interactions, and acting in accordance with social rules and conventions.” (WHO, 2007, p.174)	Disruptive behaviour Peer aggression Disregard for rules
d720 Family relationships “Creating and maintaining kinship relationships, such as with members of the nuclear family, extended family, foster and adopted family and step-relationships, more distant relationships such as second cousins, or legal guardians.” (WHO, 2007, p.176)	Attachment security Dependence Separation distress

Figure 2.2 Ternary Plot of Pooled Heritability and Shared and Nonshared Estimates



2.4 Discussion

Drawing on a systematically retrieved pooled sample of 79,044 twins, this systematic review and meta-analysis found evidence that most domains of functioning, disability, and health in psychological and developmental milestones were heritable in infancy and had moderate to high nonshared estimates. Contrary to evidence in older ages (Polderman et al., 2015), shared environment estimates were high across several important domains of infant development.

2.4.1 Heritability

Consistent with evidence in older samples (Polderman et al., 2015), all meta-analysed categories had heritability estimates with 95% CIs above 0 in infancy, apart from ‘sleep functions’ and ‘language functions’. Estimates were high (≥ 40%) for important areas of development: ‘psychomotor functions’, ‘attention functions’, ‘emotional functions’, ‘family relationships’ (which included attachment and dependency) and ‘complex interpersonal interactions’ (which encompassed behavioural problems), suggesting that phenotypes in these categories may be particularly suitable candidates for gene mapping.

High heritability in infancy of ‘attention functions’ is consistent with the high heritability of Attention Deficit Hyperactivity Disorder (ADHD) and ADHD traits in older samples (Faraone & Larsson, 2019). In accordance with the very high heritability of autism (Tick et al., 2016), a neurodevelopmental condition involving differences in social interaction, ‘social cues in relationships’ and ‘regulating behaviours within interactions’ were among the most heritable of the subcategories. Absence of evidence that infant language was heritable is consistent with evidence that the heritability of cognition, including language, is low in early development, increasing with age (Davis et al., 2009a; Plomin et al., 1997).

The higher heritability of parent-rated than observer-rated phenotypes may be driven by contrast effects in parental reports of their DZ twin children’s phenotypes, exaggerating DZ differences, or by assimilation effects in parental reports of their MZ twin children, exaggerating MZ similarities (Neale & Stevenson, 1989). Correlated rater bias that inflated both MZ and DZ twin similarity equally would lead to inflated shared environment estimates. Without raw data from individual studies, it was not possible to test this by examining variance-covariance structures, which can uncover evidence of contrast and assimilation effects. Overall, our results suggest that individual differences in growth, motor, cognitive, and emotional development can be linked to genetic factors as early as the first two years.

2.4.2 Shared Environment

Contrary to evidence in older age groups (Polderman et al., 2015), shared environment estimates had CIs above 0 in several domains and were high for ‘language functions’, ‘sleep functions’, ‘growth maintenance functions’ and ‘basic cognitive functions’, reflecting a broader trend noted in the literature that shared environmental effects on language and cognition appear to be strongest in early development (Davis et al., 2009a; Plomin et al., 1997). This may have important implications for obesity prevention and efforts to promote intellectual outcomes, which are among the most robust predictors of health and longevity (Deary et al., 2010). Shared environment estimates had CIs overlapping with 0 for ‘psychomotor functions’, ‘attention functions’, ‘emotional functions’, ‘basic interpersonal interactions’, and ‘family relationships’. This is consistent with pooled findings in older age groups (Polderman et al., 2015) and evidence that shared environments do not tend to contribute as much to similarity between siblings as genetics and do not contribute as much to differences between siblings as nonshared environments (Plomin, 2011; Plomin et al., 2016).

2.4.3 Nonshared Environment

Nonshared environment estimates had 95% CIs above 0 for all phenotypic categories and were high for ‘emotional functions’, ‘attention functions’, ‘family relationships’, and ‘basic interpersonal interactions’. Higher nonshared estimates for observer-ratings than parent ratings are consistent with wider research (Flom et al., 2018) and may reflect the importance of each twin’s unique experiences in the expression of phenotypes specifically when rated by observers. Alternatively, given that nonshared estimates also include measurement error, higher observer-rated estimates may reflect increased error in observational measurement.

2.4.4 Limitations

This study has several limitations. Given that research designs all have limitations and biases, establishing robust evidence ideally involves triangulation of methods. However, the classical twin design is currently the only quantitative genetic method that has produced data from enough independent samples to conduct adequately powered meta-analyses across a comprehensive range of infant traits. The generalizability of twin findings may be limited by the fact that some infant phenotypes (e.g., language and birth weight) develop differently in twins compared to singletons (MacGillivray et al., 1988; Ronalds et al., 2005). However, given that our aim was to examine individual differences rather than how and why groups differed, mean differences between twins and singletons may not indicate issues with generalizability.

Although the twin method can be used to examine genotype-environment correlation or interaction, we did not synthesize findings on these outcomes. In twin modelling, ignored interaction between genotype and shared environment is estimated as heritability and ignored interaction between genotype and nonshared environment will be estimated as nonshared environment, potentially contributing to biased estimates (Eaves & Rao, 1984).

Interpretation of the pooled estimates was limited by the wide CIs for some of the estimates in this study. In meta-analysis, CIs depend on the precision of included studies, which are influenced by sample size and, in the case of twin modelling, the ratio of MZ to DZ pairs and the relative contribution of each parameter (Verhulst, 2017). Furthermore, for any given sample size there is more power to estimate e^2 than h^2 and c^2, which may explain the narrower intervals around e^2. CIs also depend on the number of samples included in a meta-analysis; while adding studies can improve precision it can also increase heterogeneity, decreasing precision. Crucially, in multilevel modelling, CIs are also dependent on the degree of between-study heterogeneity. High variability in estimates across studies tends to widen their respective CIs, and ignoring such heterogeneity tends to overestimate precision. The heterogeneity observed in this study was generally high, and so CIs were comparatively wide. Comparison of estimates from *ICF-CY* categories and sub-categories was also limited by the fact that CIs were often wide and overlapping between categories. Additional moderation analysis would be of interest for future research, formally testing whether ACE estimates vary between phenotypic categories and whether estimates differ when observational measures are used, versus when a parent is the rater.

A downside of the comprehensive approach taken in this study is that it may have increased between-study heterogeneity. I attempted to reduce this in the narrower subcategory and rater analyses. However, between-study heterogeneity was substantial in all categories and subcategories, suggesting that between-study differences likely created considerable noise in our estimates. Understanding and reducing heterogeneity will be important for future research to provide more precise twin estimates in infancy. Possible publication bias was also detected across multiple outcome domains. The impact of this on the estimates is difficult to rigorously assess.

Although individuals from Western, educated, industrialized, rich and democratic societies represent approximately 12% of the world's population, twins from these areas of the world constituted more than 80% of the present sample. Infants in Africa, Asia, and South America combined represented approximately 16% of the sample, highlighting a need for twin research on infants in these continents.

There was an imbalance in the amount of research conducted on the synthesized categories; for example, far more was conducted on the anthropometric phenotypes included in ‘growth maintenance functions’ (which included data from 24 of the 52 included samples) than other domains. Important areas in which research was lacking included nonsocial autistic traits and dysregulation, eating behaviour, memory, higher level cognitive functions and brain structure.

2.4.5 Conclusion

To my knowledge, this systematic review and meta-analysis is the first study to synthesize the large and comprehensive infant twin literature on psychological traits and developmental milestones, offering insight into the possible earliest manifestations of genetic and environmental influences. This has the potential to improve public perceptions on nature and nurture by, for example, dispelling widely held beliefs that infants are shaped entirely by their environments, or that family history entirely predetermines child health, beliefs that may place undue pressure on parents. For researchers, these results offer a guide for future gene discovery research and efforts to uncover the causes of variation in infant traits. For clinicians, they provide an indication of how much family history and environmental conditions may predict infant outcomes, including outcomes that may be early markers of subsequent healthy or pathological development.

2.5. References

Austerberry, C., Mateen, M., Fearon, P., & Ronald, A. (2022). Heritability of psychological traits and developmental milestones in infancy: A systematic review and meta-analysis. *JAMA Network Open, 5*(8), e2227887-e2227887. https://doi.org/10.1001/jamanetworkopen.2022.27887

Cheung, M. W. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology, 5*(1521). https://doi.org/https://doi.org/10.3389/fpsyg.2014.01521

Darling, J. C., Bamidis, P. D., Burberry, J., & Rudolf, M. C. J. (2020). The first thousand days: Early, integrated and evidence-based approaches to improving child health: Coming to a population near you? *Archives of Disease in Childhood, 105*(9), 837–841. https://doi.org/http://dx.doi.org/10.1136/archdischild-2019-316929

Davis, O. S., Haworth, C. M., & Plomin, R. (2009). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science, 20*(10), 1301-1308. https://doi.org/10.1111/j.1467-9280.2009.02433.x

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Eaves, L. J., & Rao, D. C. (1984). The resolution of genotype × environment interaction in segregation analysis of nuclear families. *Genetic Epidemiology, 1*(3), 215-228. https://doi.org/10.1002/gepi.1370010302

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/https://doi.org/10.1136/bmj.315.7109.629

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

Flom, M., Wang, M., Uccello, K. J., & Saudino, K. J. (2018). Parent- and observer-rated positive affect in early childhood: Genetic overlap and environmental specificity. *Behavior Genetics, 48*(6), 432-439. https://doi.org/10.1007/s10519-018-9924-0

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry, 46*(3), 289. https://doi.org/https://doi.org/10.1111/j.1465-7295.2008.00163.x

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560. https://doi.org/https://doi.org/10.1136/bmj.327.7414.557

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*.

Kumsta, R., Kreppner, J., Kennedy, M., Knights, N., Rutter, M., & Sonuga-Barke, E. (2015). Psychological consequences of early global deprivation: An overview of findings from the English & Romanian Adoptees study. *European Psychologist, 20*(2), 138. https://doi.org/10.1027/1016-9040/a000227

Lejarraga, H. (2012). Growth in infancy and childhood: A pediatric approach. In N. Cameron & B. Bogin (Eds.), *Human Growth and Development* (Second ed., pp. 23–56). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-383882-7.00002-7

MacGillivray, I., Campbell, D. M., & Thompson, B. (1988). *Twinning and twins*. Wiley.

Neale, M. C., & Stevenson, J. (1989). Rater bias in the EASI temperament scales: A twin study. *Journal of Personality and Social Psychology, 56*(3), 446-455. https://doi.org/10.1037//0022-3514.56.3.446

Nelson, C. A., 3rd, Zeanah, C. H., & Fox, N. A. (2019). How early experience shapes human development: The case of psychosocial deprivation. *Neural Plasticity, 2019*. https://doi.org/https://doi.org/10.1155/2019/1676285

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal, 372*. https://doi.org/https://doi.org/10.1136/bmj.n71

Papageorgiou, K. A., & Ronald, A. (2017). The genetic basis of psychological traits in infancy: Implications for understanding the causes of developmental psychopathology. In L. C. Centifanti & D. M. Williams (Eds.), *The Wiley handbook of developmental psychopathology* (pp. 235–258). Wiley Blackwell. https://doi.org/https://doi.org/10.1002/9781118554470.ch11

Pappa, I., Szekely, E., Mileva-Seitz, V. R., Luijk, M. P., Bakermans-Kranenburg, M. J., van, I. M. H., & Tiemeier, H. (2015). Beyond the usual suspects: A multidimensional genetic exploration of infant attachment disorganization and security. *Attachment and Human Development, 17*(3), 288–301. https://doi.org/https://doi.org/10.1080/14616734.2015.1037316

Plomin, R. (2011). Commentary: Why are children in the same family so different? Non-shared environment three decades later. *International Journal of Epidemiology, 40*(3), 582-592. https://doi.org/10.1093/ije/dyq144

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science, 11*(1), 3-23. https://doi.org/10.1177/1745691615617439

Plomin, R., Fulker, D. W., Corley, R. P., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 Years: A parent-offspring adoption study. *Psychological Science, 8*(6), 442-447. https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics, 47*(7), 702-709. https://doi.org/10.1038/ng.3285

R Core Team. (2020). *A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ronald, A., Happe, F., & Plomin, R. (2008). A twin study investigating the genetic and environmental aetiologies of parent, teacher and child ratings of autistic-like traits and their overlap. *European Child and Adolescent Psychiatry, 17*(8), 473–483. https://doi.org/https://doi.org/10.1007/s00787-008-0689-5

Ronalds, G. A., De Stavola, B. L., & Leon, D. A. (2005). The cognitive cost of being a twin: Evidence from comparisons within families in the Aberdeen children of the 1950s cohort study. *British Medical Journal, 331*(7528), 1306. https://doi.org/https://doi.org/10.1136/bmj.38633.594387.3A

Rutter, M. (1998, May). Developmental catch-up, and deficit, following adoption after severe global early privation. *The Journal of Child Psychology and Psychiatry, 39*(4), 465–476. https://www.ncbi.nlm.nih.gov/pubmed/9599775

Saudino, K. J., Ronald, A., & Plomin, R. (2005). The etiology of behavior problems in 7-year-old twins: Substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. *Journal of Abnormal Child Psychology, 33*(1), 113–130. https://doi.org/https://doi.org/10.1007/s10802-005-0939-7

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. Meta-analysis Of Observational Studies in Epidemiology

(MOOSE) group. *JAMA, 283*(15), 2008-2012. https://doi.org/10.1001/jama.283.15.2008

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

Verhulst, B. (2017). A power calculator for the classical twin design. *Behavior Genetics, 47*(2), 255-261. https://doi.org/10.1007/s10519-016-9828-9

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

World Health Organization. (2007). *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY*. World Health Organization.

Chapter 3: Early Manifestations of Academic Performance: Evidence That Genetic Effects on Later Academic Test Performance are Mediated Through Language in Early Childhood

A version of this study was published in *Child Development* (Austerberry, Fearon, et al., 2022).

3.1 Introduction

Intellectual and academic abilities are among the strongest predictors of lifelong success, health and longevity (Deary et al., 2010; Hummer & Hernandez, 2013; Kosik et al., 2018), and both are substantially heritable (Bouchard & McGue, 1981; Kovas et al., 2013). However, the evidence base is small regarding the earliest manifestations of genetic influences on intellectual and academic abilities. These early manifestations are important because they may be in the causal chain from genetic influences to later intellectual and academic performance and because they are also likely to have an indirect influence on the development of intellectual and academic abilities through interplay with caregiving and learning environments. I used a parent-offspring adoption design to examine two likely candidate early manifestations: Executive function (EF) and language performance in early to middle childhood. These results are the first to document whether early EF or language have a mediating role, linking genetic influences to later intellectual and academic performance in middle childhood and possibly also in adulthood. By identifying which of these, EF or language, serves as a principal manifestation of genetic influences, these results pave the way for investigations into how children's interactions with parents and teachers from early childhood onwards amplify or diminish these favourable outcomes.

Intellectual and academic abilities are powerful predictors of psychological wellbeing, health, longevity, years of education, income, and employment status (Deary et al., 2010; Hummer & Hernandez, 2013; Kosik et al., 2018). Conversely, lower intellectual performance is associated with all-cause mortality and clinically important increases in the severity of psychopathology (Deary et al., 2010; Kosik et al., 2018; Yew & O'Kearney, 2013). Furthermore, there is evidence to suggest that academic performance in adolescence may have a negative causal connection with internalizing and externalizing problems in emerging adulthood (Wolchik et al., 2016).

Consequently, promotion of intellectual and academic performance in childhood may have broad effects across development, including improving educational, occupational and health outcomes, and diminishing the likelihood of some psychiatric problems. As a result, research aimed at understanding the processes involved in the early development of intellectual and academic abilities is crucial and may help uncover mechanisms that can be modified, not only to promote intellectual and academic development, but also to promote a wide range of positive life outcomes and reduce the risk of psychopathology.

Intellectual and academic abilities have been reported to be highly heritable, especially as children get older, rising from 20–60% in childhood and adolescence to 50–80% in adulthood (Bouchard & McGue, 1981; C. M. Haworth et al., 2010; Kovas et al., 2013). Consequently, some have argued that environmental factors must play only a minor role in intellectual development (Plomin, 2018). However, twin and adoption studies provide evidence that environmental factors can have notable main effects and moderating effects on intellectual outcomes (Capron & Duyme, 1989; Kendler et al., 2015; Neiss & Rowe, 2000; Tucker-Drob & Bates, 2015). There is also evidence from the recent surge of literature using measured genotypes to examine genetic nurture (Bates et al., 2018; Kong et al., 2018; Wertz et al., 2020)—including studies that have combined polygenic scores with the adoption design (Cheesman et al., 2020; Domingue & Fletcher, 2020)—suggesting that parents influence children's academic outcomes not only through direct genetic transmission but also through environmentally mediated pathways. An additional, unheralded, mechanism is that the environment may have an amplifying effect on genetic influences, through evocative gene-environment correlation (*r*GE). This occurs when an individual's genetically influenced characteristics systematically evoke responses from their environment that, in turn, enhance or 'canalize' genetic influences (Scarr & McCartney, 1983). As these evoked environmental conditions correlate with genetic influences, their influences could be masked by estimates of genetic main effects. Dickens and Flynn (2001) explore in detail the possibility that this process of amplification operates in the context of cognitive abilities across generations to account for rising levels of intelligence in successive cohorts of children and adults. While there is some evidence from phenotypic, twin and polygenic score research of evocative *r*GE in infant and early childhood cognitive development (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012; Wertz et al., 2020), the evidence base is small and the Dickens and Flynn hypothesis has never been robustly tested across the span of development within a generation. For these environmental amplification effects to be examined in detail, it is important to know at which developmental periods they may exert their influence on intellectual and academic outcomes. For influences occurring early in development it is crucial to identify the earliest manifestations of genetic advantage because these are likely the features that elicit the favourable environmental responses that amplify genetic effects.

In spite of the great importance of identifying the early manifestations of genetic influences on lifespan intellectual and academic abilities, the evidence base is small with regards to what these early manifestations might be. There is some indication that childhood scholastic performance from 6–7 years old onwards may be an early indicator of genetic advantage for academic performance in adulthood. For example, higher genome-wide polygenic scores of total years of education achieved by adulthood (EduYears) predict stronger reading and math test performance at 6, 7, 12 and 16 years (Allegrini et al., 2019; Armstrong-Carter et al., 2020; Belsky et al., 2016; Selzam et al., 2017). This is supported by evidence that adopted children's math and reading performance at age 7 years is partially predicted by their birth parents' reading and math test performance (Borriello et al., 2020; Cioffi et al., 2021). These associations are not confounded by direct caregiving effects because adopted children and birth parents share genes, but birth parents do not provide the postnatal rearing environment. Furthermore, as the birth parent outcomes were measured in adulthood, the observed phenotypic associations between birth parents and children are akin to an "instant longitudinal study" from childhood to adulthood (Plomin, 1986) because, although these studies do not include longitudinal data from childhood to adulthood, they identify genetic factors accounting for the association between academic test performance in childhood and academic performance of biological relatives in adulthood. It remains less well understood whether there are earlier markers of genetic effects on lifespan intellectual performance than academic test performance from age 6–7 years onwards. There is mixed evidence from one longitudinal study (the Dunedin

Study): Children in the sample with higher EduYears polygenic scores began talking earlier, based on parent ratings of developmental milestones at 3 years old, but did not score any better in the Peabody Picture Vocabulary Test at 3 years old (Belsky et al., 2016). However, from 5 years old, and onwards, children in the study with higher EduYears polygenic scores scored higher on tests of intelligence. Additionally, there is some evidence from adoption studies that birth parent intellectual performance in adulthood predicts adopted children's executive function (EF), verbal and non-verbal intelligence from 1–3 years old (Leve, DeGarmo, et al., 2013; Plomin et al., 1997), indicating that these early abilities may be markers of genetic effects on adult intellectual performance. This is consistent with evidence, firstly, that EF and language in early childhood are partially heritable, including from as early as two years old, at which point, as seen in **Chapter 2** of this thesis, they have low heritability estimates (h^2_{pooled}, 24% and 34%, respectively) (Austerberry, Mateen, et al., 2022), and throughout early and middle childhood, by which point the heritability of EF and language appear to be approximately 60% (Davis et al., 2009a; Polderman et al., 2007). Secondly, these findings are consistent with evidence that early childhood EF and language predict later intellectual performance (Duncan et al., 2007; McClelland et al., 2013; Yu et al., 2018). For example, there is longitudinal evidence from six studies that reading, language and attention at school entry robustly predict later school math and reading test performance (Duncan et al., 2007). It seems likely, based on these converging bodies of research, that EF and language in early and middle childhood are early manifestations of genetic effects on later intellectual and academic performance. However, no research has used an adoption design to combine these streams of evidence and investigate whether early (and apparently heritable) EF and language mediate genetic influences on later intellectual and academic outcomes.

Early manifestations of genetic effects on intellectual and academic performance are important to understand, firstly, because they may directly influence the development of later intellectual and academic performance. Secondly, because they likely have an indirect influence on the development of intellectual and academic abilities through interplay with caregiving and learning environments that plausibly sustain and amplify these early manifestations.

However, as there was limited evidence of what the early manifestations are and precisely when they manifested, research was not yet in a position to rigorously explore hypotheses about evocative effects of genetic influences underlying intellectual and academic development on caregiving and learning environments. I took the critical first step of identifying the very early expressions of genetic advantage in intellectual and academic performance using a longitudinal parent-offspring adoption study.

I examined the hypothesis that genetic effects linked to adult intellectual and academic performance would have their impact on child reading and maths test performance at 7 years of age through two early appearing pathways: via EF and language performance from 27 months to 6 years. This hypothesis was addressed in two steps: First, I examined at which age, or ages (between the 27 months and 6 years) there was evidence of genetic effects on EF and language. Second, if the first set of analyses confirmed my expectation that there would be evidence of genetic effects on early EF and language, I tested for mediation of genetic effects on reading and maths test performance at 7 years old via each of these pathways. My expectation was that early EF and language would mediate genetic effects on intellectual and academic performance, indicating that they are early manifestations of genetic influences on these assets. I used birth mother general intellectual and academic performance—captured using a latent composite of intelligence, reading and math test performance—as a proxy for genetic influences. As adopted children and their birth mothers share genes but birth mothers do not provide the postnatal rearing environment, the adoption design eliminates the influence of birth mothers on the postnatal environment. Phenotypic associations between adopted children and their birth mothers would thus be taken to imply genetic effects. However, correlations between birth mothers and their adopted offspring can represent a combination of genetic and prenatal environment effects. Birth fathers, who play an equal role to mothers in contributing to the child's genotype, provide an estimate of genetic effects that is not confounded by prenatal effects. Consequently, I used a smaller subsample of birth fathers for replications of the birth mother analyses. Although birth parents tend to correlate on measures of intelligence (Bouchard & McGue, 1981) and the birth father replications could only be considered quasi-independent rather than fully independent replications, broadly speaking, they provided convergent evidence regarding the effects of genetic influences (as opposed to the prenatal environment) on children's intellectual and academic outcomes.

3.2 Method

3.2.1 Participants

Participants were drawn from the Early Growth and Development Study (EGDS), a US-based, longitudinal, prospective adoption study of 561 linked sets of adopted children and their birth mothers (*n* = 554), birth fathers (*n* = 210) and adoptive parents (562 adoptive fathers and 569 adoptive mothers; numbers do not sum to 561 adoptive mothers and fathers because the sample includes 41 same-sex parent families and 15 additional adoptive parents who entered the family after the original couple adopted the child) (Leve et al., 2019; Leve, Neiderhiser, et al., 2013a). EGDS data were collected in two cohorts, recruited through 45 adoption agencies in 15 states across the US (Leve et al., 2019): The first, a sample of 361 adopted children and their birth and adoptive families and, the second, a sample of 200. While some of the variables used in the analysis were collected in both cohorts, others were only collected in one cohort. For a breakdown of the number of participants by each variable, see Figure 3.1 and Table 3.1. The variables used in the present analyses were collected in 2003–2013 (cohort I) and in 2007–2017 (cohort II).

Mean age of children at adoption placement was 5.58 days postpartum (*SD* = 12.4; median = 2; range = 0–91). Over half of the adopted children were male (57%). The children were 56% non-Latinx White, 19% multiracial, 13% Black/African American, 11% Latinx and <1% "other" (including Asian, American Indian, and unknown ethnicity). Adoptive parents were typically non-Latinx White (adoptive parent 1: 92%; adoptive parent 2: 90%). The remainder were Black/African American (adoptive parent 1: 4%; adoptive parent 2: 5%), Hispanic or Latinx (adoptive parent 1: 2%; adoptive parent 2: 2%), multiracial (adoptive parent 1: 1%; adoptive parent 2: 1%), and "other" (adoptive parent 1: 1%; adoptive parent 2: 2%). At the time of adoption, adoptive parents were typically in their mid- to late-thirties (adoptive parent 1: *M* = 37.4, *SD* = 5.6; adoptive parent 2: *M* = 38.3, *SD* = 5.8), married or cohabiting (adoptive parent 1: 98%, adoptive parent 2: 100%), college educated and with a combined median income above $100,000. Birth parents were typically non-Latinx White (mothers: 70%; fathers: 70%). The remainder were Black/African American (mothers: 13%; fathers: 12%), Hispanic or Latinx (mothers: 7%; fathers: 10%), multiracial (mothers: 5%; fathers: 5%), and "other" (mothers: 5%; fathers: 4%). At the time of adoption, birth parents were typically in their mid-twenties (mothers: $M = 24.4$, $SD = 6.0$; fathers: $M = 26.1$, $SD = 7.8$), married or cohabiting (mothers: 6.1%; fathers: 14.0%), had less than a college education, and median household incomes below $25,000. There is no evidence of selective placement in EGDS (Leve et al., 2019). Additional information about the recruitment, composition and representativeness of the sample is reported elsewhere (Leve et al., 2019). EGDS assessments are ongoing and occurred in intervals of 9 months to 2 years. I used pre-existing data collected from birth parents at 18 months, 4.5 and 7 years postpartum and data collected from adoptive parents and adoptees when the children were aged 27 months, 4.5 years, 6 years, and 7 years.

3.2.2 Ethics

Ethical approval was obtained from institutional review boards at the University of Oregon (Protocol number: 0304201400) and The Pennsylvania State University (Submission ID: CR00007591). Informed consent was obtained from all adult participants ahead of research participation and assent was obtained from children at age 7 years.

3.2.3 Measures

Using structural equation models, incorporating confirmatory factor analysis (CFA), I created the latent variables (displayed in Figure 3.1) across each of the domains outlined below. Prior to hypothesis testing, I ran longitudinal measurement models, assessing the fit of individual domains across all timepoints. Model fit was good in all these models, supporting the use of latent variables.

Birth Parent General Intellectual Performance. As displayed in Figure 3.1, I created a latent variable of birth parent general intellectual performance, with the indicators of intelligence, reading and math test performance listed below, as a proxy for genetic influences on children. Latent measurement drawing on a diverse range of indicators was justified by the internal consistency (birth mother $\alpha_R = .84$; birth father $\alpha_R = .85$) and bivariate correlations among measures of birth parent intelligence and academic test performance in the EGDS sample (Table 3.1), and the ‘generalist genes’ literature which reports that approximately a third of the genetic variance of reading and math performance is in common with general intelligence (*g*) (Davis et al., 2009a; Plomin & Kovas, 2005).

Wechsler Adult Intelligence Scale-III. EGDS administered the 28-item Information subtest (Wechsler, 1997) to birth parents at 18 months postpartum. This subtest loads onto the verbal comprehension index of the full measure and is considered to be a representative measure of *g* (*g* loading = .79) (Kaufman & Lichtenberger, 1999). I used standardized scores, based on age.

Woodcock-Johnson Tests of Achievement III. At 4.5 or 7 years postpartum, EGDS administered birth parents four subtests: (1) 76-item Letter-Word Identification, measuring reading decoding; (2) 32-item Word Attack, capturing decoding and phonetic coding; (3) 98-item Reading Fluency, measuring reading speed and semantic processing speed; (4) 160-item Math Fluency, indexing math and numerical performance (Woodcock et al., 2001). I used T-scores, standardized to have a mean of 50 and standard deviation of 10.

Child Executive Function. As displayed in Figure 3.1, I created latent measures of executive function (EF) at 27 months ($r = .15$), 4.5 years ($\alpha_R = .38$) and 6 years ($\alpha_R = .41$), with and the EF indicators listed below. Use of latent variables was justified by the model fit (RMSEA =.03, SRMR = .05) and longitudinal stability of EF from 27 months to 4.5 years (β = .74, 95% CI [.41, 1.07], $p < .001$) and from 4.5 years to 6 years (β = .70, 95% CI [.45, .95], $p < .001$) in the longitudinal measurement model, as well as by the concurrent and longitudinal correlations between these EF measures, in the EGDS sample (Table 3.1) and wider literature (Montgomery & Koeltzow, 2010).

Stroop Task. At 27 months, EGDS administered the fruits-animals Stroop, modified by the EGDS team based on Kochanska et al. (2000). There were six trials, each scored on a scale from 1 to 3 (1 = incorrect on item and size; 2 = correct item, wrong size; 3 = correct item and size). The trials had strong internal consistency (α = .85) and were averaged to form a scale score. At 4.5 and 6 years, EGDS administered the 16-trial day-night Stroop (Gerstadt et al., 1994), which has robust construct validity and internal and test-retest reliability (Montgomery & Koeltzow, 2010). Each trial had one point for a correct answer. Trials had strong internal consistency (α = .85) and were summed, resulting in a score between 0 and 16.

Gift Delay Task. At 27 months, children participated in a gift delay task similar to the one described by Kochanska et al. (1996). EGDS raters coded the videotaped task based on how often the child (1) peeked, (2) touched the gift, and (3) used distraction strategies. In line with Leve, DeGarmo, et al. (2013), I averaged the 3 items to form a total score of inhibitory control, with higher scores indicating higher inhibition (α = .54; *r* = .08, .32, and .46 among items).

Guessing Game. At 4.5 years old, children completed a task adapted from the Goldsmith and Rothbart (1999) laboratory assessment of temperament (Lab TAB) to measure their inhibitory control when told not to turn around or peek at hidden toys. The task was coded by EGDS raters from 1 (Not at all) to 5 (Continually) on: "How often did the child keep their back turned around when asked to?".

Forbidden Gift. EGDS measured inhibitory control in the 4.5-year-olds using a forbidden gift task modified from the Lab TAB (Goldsmith & Rothbart, 1999), which was videotaped and coded by EGDS raters from 1 (Very True) to 3 (Not true) on whether: "The child asked for the gift".

Dinky Toys. This inhibitory control task modified from the Lab TAB (Goldsmith & Rothbart, 1999) by EGDS researchers involved the 4.5- and 6-year-olds being asked to comply with rules about how to interact with a box of toys. EGDS raters coded the task on: "The degree to which the child follows or violates instructions" from 1 (violates rules) to 3 (follows all instructions).

Go-NoGo. At 6 years, EGDS administered a Go-NoGo task (Nosek & Banaji, 2001). In this 84-trial version, trials were divided into two blocks, the first of which contained only Go trials (when the child should press a button) and the second an equal combination of Go trials and NoGo trials (in which children are expected to inhibit their prepotent response by refraining from pressing a button). I used the percentage of correct responses in the second block to both Go and NoGo stimuli as a measure of selective attention and inhibition.

Child Language Performance. I created latent variables at 27 months (*r* = .74), 4.5 years (α_R = .62), and 6 years (α_R = .76) with the indicators of language listed below. My decision to use latent variables combining these indicators was guided by the model fit (RMSEA = 0.06; SRMR = 0.07) and longitudinal stability from 27 months to 4.5 years (β = .47, 95% CI [.35, .59], *p* < .001) and from 4.5 years to 6 years (β = .76, 95% CI [.66, .87], *p* < .001) in the longitudinal measurement model, as well as by the concurrent and bivariate correlations between these measures in the EGDS sample (Table 3.1) and wider literature (Sim et al., 2019).

Language Development scale (LDS). Adoptive parents separately completed a measure of child language development at 27 months, based on the number of words that the child is reported to use spontaneously from a list of 310 items (Achenbach & Rescorla, 2000). Reports from both parents were used as indicators in the language latent variable at 27 months. Using nationally standardized normed scores, EGDS converted raw scores to percentiles that reflected the child's language performance relative to same-age peers (Achenbach & Rescorla, 2000). LDS scores have moderate to high correlations (*r* = .66 – .87) with scores on standardized vocabulary tests (Klee et al., 1998; Rescorla & Alley, 2001) and are reported to have the best predictive validity performance of the language screening tools (Sim et al., 2019).

Test of Preschool Early Literacy (TOPEL). EGDS administered three subscales to 4.5-year-olds (Lonigan et al., 2007): (1) 36-item Print Knowledge, measuring knowledge of the alphabet, written language conventions and written form; (2) 35-item Definitional Vocabulary, assessing definitional and single-word oral vocabulary; (3) 27-item Phonological Awareness, measuring word elision and blending. I used standard scores, derived from the distribution of the raw scores. The TOPEL has high internal consistency (α = .86 – .96) and test-retest reliability (*r* = .81 – 89), moderate predictive validity (*r* = .40 – .62) and moderate to high concurrent validity (*r* = .59 – .77) (Lonigan et al., 2011).

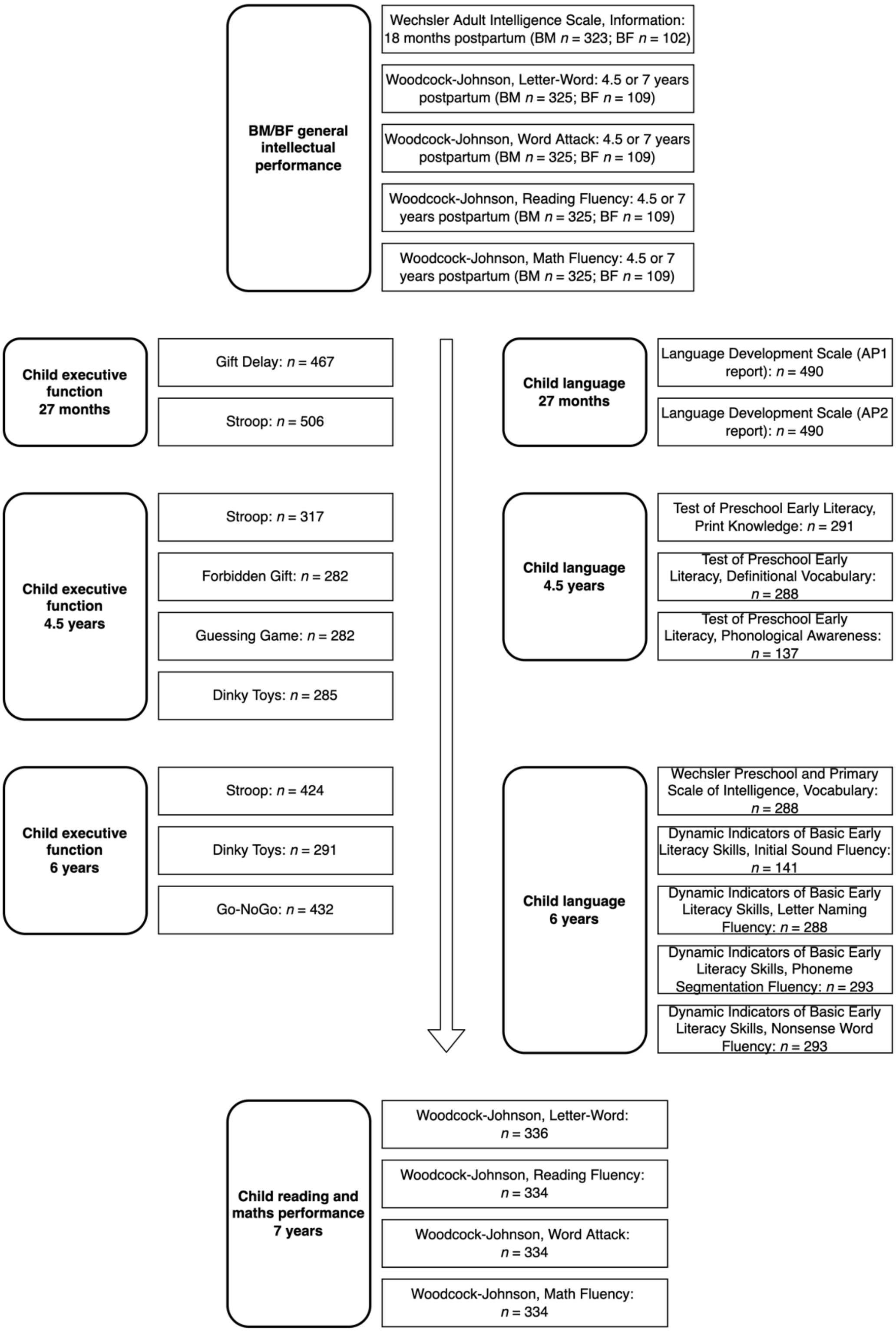
Wechsler Preschool and Primary Scale of Intelligence III (WPPSI). EGDS administered the vocabulary assessment to 6-year-olds, measuring learning, comprehension and verbal expression of vocabulary (Wechsler, 2002). I used raw scores from the 50-item measure which had been converted by EGDS to standardized scores from 1 to 19, based on the responder's age.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS). EGDS administered four sets of procedures and assessments to 6-year-olds: (1) 16-item Initial Sound Fluency (ISF), measuring phonemic awareness; (2) Letter Naming Fluency (LNF), capturing proficiency in naming upper- and lower-case letters, using a list of 110 letters; (3) Phoneme Segmentation Fluency (PSF), assessing proficiency in fluently segmenting three- and four-phoneme words into their individual phonemes, using a list of 24 words; (4) Nonsense Word Fluency (NWF), testing understanding of the alphabetic principle, including letter-sound correspondence, using a list of 50 nonsense words (Good & Kaminski, 2002). ISF and LNF have good test-retest reliability (*r* = .88 – .93) and robustly predict later reading performance (Kaminski & Good, 1996). Raw scores, which represent the number of items a child has answered correctly in one minute, were converted by EGDS to percentiles, reflecting language relative to same grade level peers, based on nationally standardized normed scores (Good & Kaminski, 2002).

Child Academic Test Performance. Justified by the high genetic correlations between reading and math performance in childhood (Davis et al., 2009b; Plomin & Kovas, 2005), and the internal consistency (α = .87) and bivariate correlations in the EGDS sample (Table 3.1), I created a latent variable to estimate child academic test performance at 7 years old, drawing on the same four indicators of reading and math performance that were administered to birth parents from the Woodcock-Johnson Tests of Achievement III (Woodcock et al., 2001)—see Figure 3.1.

Covariates. I included adoption openness, sex of child, and prenatal risk as covariates. I used a mean standardized composite of birth mother and adoptive parent reported adoption openness, using a 4-item measure (Ge et al., 2008), averaged across ratings provided at 9, 18, and 27 months postpartum. EGDS collected birth mother reports of maternal and pregnancy complications, labour and delivery complications and neonatal complications at 5 months postpartum and scored them based on the 76-item McNeil-Sjostrom Scale for Obstetric Complications (McNeil et al., 1994). I used a weighted total prenatal risk score created by EGDS based on work by Marceau et al. (2016).

Figure 3.1 Latent and Observed Variables Used in the Structural Equation Models



Note. Latent variables are displayed in bold and observed variables are displayed in regular text. BM = birth mother. BF = birth father. WAIS = Wechsler Adult Intelligence Scale-III. WJ = Woodcock-Johnson Tests of Achievement III. LDS = Language Development Scale. AP1 = adoptive parent 1. AP2 = adoptive parent 2. TOPEL = Tests of Preschool Early Literacy. WPPSI = Wechsler Preschool and Primary Scale of Intelligence III. DIBELS = Dynamic Indicators of Basic Early Literacy Skills.

Table 3.1 Means, Standard Deviations, Sample Size, and Bivariate Correlations Between Study Variables

Variable	*M*	*SD*	*n*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1. 27 mo. Gift Delay	1.96	0.76	467	—																															
2. 27 mo. Stroop	2.01	0.57	506	.15	—																														
3. 27 mo. LDS AP1	57.61	21.83	490	.13	.29	—																													
4. 27 mo. LDS AP2	52.69	22.79	464	.12	.23	.74	—																												
5. 4.5 yr. Stroop	9.64	5.90	317	.08	.14	.07	.14	—																											
6. 4.5 yr. Forbidden Gift	2.29	0.78	282	.11	.04	-.06	-.04	.17	—																										
7. 4.5 yr. Guessing Game	4.25	1.16	282	.21	.07	.17	.14	.17	.14	—																									
8. 4.5 yr. Dinky Toys	2.03	0.69	285	.06	.07	.04	.08	.09	.05	.22	—																								
9. 4.5 yr. TOPEL PK	106.76	13.39	291	.15	.20	.27	.23	.17	.19	.26	.15	—																							
10. 4.5 yr. TOPEL DV	104.29	9.94	288	.19	.25	.30	.29	.15	.21	.27	.24	.43	—																						
11. 4.5 yr. TOPEL PA	97.76	15.66	137	.21	.18	.23	.24	.01	.10	.20	.20	.38	.30	—																					
12. 6 yr. Stroop	14.20	3.46	424	.13	.07	.06	.02	.07	-.07	.14	.06	.06	.10	.20	—																				
13. 6 yr. Dinky Toys	2.13	0.65	291	.13	.15	.08	.09	.10	.01	.12	.29	.12	.17	.06	.08	—																			
14. 6 yr. Go NoGo	73.94	12.53	432	.09	.12	.13	.16	.20	-.02	.18	.14	.09	.20	.24	.27	.21	—																		
15. 6 yr. WPPSI	10.31	2.28	288	.10	.27	.29	.20	.07	-.02	.21	.18	.34	.46	.31	.08	.08	.13	—																	
16. 6 yr. DIBELS ISF	58.48	30.22	141	.08	.12	.18	.13	.16	.04	.16	.28	.37	.35	.14	.02	.24	.22	.24	—																
17. 6 yr. DIBELS LNF	52.18	31.68	288	.14	.09	.12	.11	.14	.08	.17	.16	.37	.26	.14	.15	.10	.09	.29	.62	—															
18. 6 yr. DIBELS PSF	33.84	27.87	293	.04	.14	.24	.19	.23	.11	.16	.18	.31	.30	.23	.17	.08	.18	.29	.26	.44	—														
19. 6 yr. DIBELS NWF	46.19	32.58	293	.03	.16	.27	.23	.14	.04	.15	.21	.47	.29	.35	.18	.13	.08	.36	.39	.56	.52	—													
20. 7 yr. W-J Letter Word	56.29	9.23	336	.09	.20	.23	.19	.14	.09	.19	.14	.49	.32	.34	.13	.08	.11	.34	.29	.38	.36	.62	—												
21. 7 yr. W-J Reading	53.91	10.80	334	.08	.21	.19	.13	.10	.15	.21	.15	.51	.35	.34	.12	.11	.11	.36	.30	.37	.34	.59	.83	—											
22. 7 yr. W-J Word Attack	55.22	7.51	334	.09	.18	.22	.18	.10	.05	.16	.15	.36	.29	.43	.13	.09	.07	.36	.24	.32	.37	.54	.82	.66	—										
23. 7 yr. W-J Math	49.66	10.15	334	.07	.12	.14	.11	.07	.11	.13	.06	.28	.19	.22	.16	-.03	.07	.29	-.03	.21	.24	.38	.56	.58	.47	—									
24. BM WAIS Information	9.56	2.59	323	.10	.25	.11	.07	.01	.10	.07	.03	.23	.18	.03	-.04	.07	-.11	.16	.22	.12	.12	.12	.23	.18	.23	.15	—								
25. BM W-J Letter Word	47.89	5.85	325	.07	.14	.09	.08	.03	.06	-.02	.09	.22	.09	.10	.02	.13	.03	.12	.22	.09	.10	.15	.25	.22	.20	.08	.57	—							
26. BM W-J Reading	49.80	7.53	325	.10	.10	.10	.08	.04	.03	.10	.02	.25	.17	.11	.04	.11	.01	.12	.15	.18	.17	.23	.21	.18	.17	.08	.49	.59	—						
27. BM W-J Word Attack	46.87	6.73	325	-.03	.17	.07	.09	.02	.02	.01	.10	.13	.08	.10	.16	.14	.03	.07	.26	.08	.05	.12	.22	.19	.20	.08	.37	.70	.49	—					
28. BM W-J Math	44.04	8.86	325	.02	.06	.01	.00	.01	.07	.06	-.06	.23	.04	.02	.06	.02	-.01	.12	.18	.18	.12	.18	.13	.12	.10	.14	.35	.46	.62	.44	—				
29. BF WAIS Information	10.65	2.88	102	-.08	.12	.01	.09	-.03	-.11	.05	.16	.06	.04	-.01	-.10	.10	-.10	.22	-.01	-.10	.01	-.05	-.10	-.17	-.09	.05	.42	.37	.26	.17	.24	—			
30. BF W-J Letter Word	47.29	7.67	109	.08	.01	.12	.08	-.02	.22	.15	.24	.18	.05	.42	.08	.14	.08	.22	.25	.25	.18	.28	.27	.30	.24	.43	.45	.33	.34	.25	.22	.57	—		
31. BF W-J Reading	46.60	7.80	109	-.02	-.08	.07	.11	.06	.11	.23	.20	.11	.06	.15	.15	.06	.14	.22	.22	.20	.17	.15	.06	.03	.23	.14	.30	.22	.20	.14	.18	.54	.61	—	
32. BF W-J Word Attack	46.94	8.40	109	-.01	.01	.10	.11	.03	.14	.02	.13	.17	.01	.25	.16	.01	.07	.13	.16	.21	.24	.28	.27	.25	.21	.37	.40	.35	.29	.17	.26	.46	.76	.66	—
33. BF W-J Math	41.95	10.01	109	.09	.10	.03	.04	-.01	.05	.17	.22	.01	.03	.20	.00	.04	.08	.19	.27	.17	.09	.01	.05	.03	.19	.21	.33	.26	.24	.22	.28	.32	.51	.63	.55

█ $p < .001$
▓ $p < .01$
▒ $p < .05$
░ $p < .1$
□ $p > .1$

Note. LDS = Language Development Scale, AP1 = adoptive parent 1 report; AP2 = adoptive parent 2 report; TOPEL = Test of Preschool Early Literacy PK = Print Knowledge; DV = Definitional Vocabulary; PA = Phonological Awareness; WPPSI = Wechsler Preschool and Primary Scale of Intelligence III; DIBELS = Dynamic Indicators of Basic Early Literacy Skills; ISF = Initial Sound Fluency; LNF = Letter Naming Fluency; PSF = Phoneme Segmentation Fluency; NWF = Nonsense Word Fluency; W-J = Woodcock-Johnson Test of Achievement III; BM = birth mother; BF = birth father; WAIS = Wechsler Adult Intelligence Scale-III

3.2.4 Data Analysis

I conducted the primary analyses using birth mother and child data only and used data from a smaller subsample of birth fathers to carry out a quasi-independent replication. Although the birth father sample is the largest ever recruited in a prospective parent-offspring adoption study, it has reduced statistical power compared to the birth mother analyses. Thus, I anticipated that comparisons between results for birth mothers and birth fathers would focus on the magnitude of the path coefficients rather than *p* values or confidence intervals.

I tested the hypothesis in two steps, in the lavaan package (Rosseel, 2012) in R 4.0.0, using structural equation modelling, which combines a measurement model (also known as CFA) with a structural model testing the proposed causal relations. First, I constructed longitudinal models examining: (1) Whether EF and language were stable across 27 months, 4.5 years, and 6 years, and predicted academic test performance at 7 years; and (2) whether there were genetic effects on child EF, language, and academic test performance. Second, if the models were consistent with mediation of genetic effects on academic test performance at 7 years through early EF or language, I ran mediation models examining the indirect effects of birth parent intellectual performance on child academic test performance at 7 years. I included the covariates in all of the models and used bootstrapping with 5000 repetitions to test the indirect effect in the mediation models (Bollen & Stine, 1990). Based on recommendations by Hu and Bentler (1999), I used a combination rule, according to which model fit was considered adequate if SRMR < .09 and RMSEA < .06.

Variable sample sizes are reported in Table 3.1 and Figure 3.1. The primary source of missing data in child EF models using birth mother data was child EF measures at 4.5 years. In child language models using birth mother data it was the DIBELS ISF subscale. In birth father and child models, it was missing information on birth father intellectual performance. The data used in the analyses were not missing completely at random [Little's MCAR χ^2 (4598) = 4884.36, $p < .01$]. I ran an attrition analysis using the Missing Value Analysis function in SPSS, which creates an indicator variable identifying variables that contain missing values. This indicator value is then used to compare group means among different variables in the dataset, using the t-test procedure. The attrition analysis revealed that the patterns of missingness for the majority (69%) of study variables were related to the observed values of one or more other variables in the dataset. This analysis ruled out the possibility that the data were missing completely at random (MCAR), which occurs when the probability of being missing is the same for all cases and there is no systematic association between the missingness of the data and any other values, observed or missing. It was not possible to rule out the possibility that the data were missing not at random (MNAR), which is when the missingness of the data is systematically related to unobserved data. However, the associations found in the attrition analysis are consistent with the data being missing at random (MAR), which occurs when the missingness of a variable is systematically related to the observed but not the unobserved data. I handled missing data in the analyses using full information maximum likelihood (FIML), which is suitable for data that is MAR, and is of comparable performance to multiple imputation (Allison, 2003).

3.2.5 Sensitivity Analyses

I conducted sensitivity analyses to examine: (1) the impact of the removal of earlier time points on associations between birth parent general intellectual performance and child EF, language and academic test performance; (2) whether the age at which birth parents were administered measures of intellectual performance was associated with their intellectual performance and, if so, whether birth parent age confounded the associations between birth parents and children; and (3) if the indirect association between birth parent intellectual performance and academic test performance at 7 years, via children's earlier language still held when the mediation models were recomputed using only the math subscale of the academic test performance measure at 7 years. The third sensitivity analysis was conducted as a robustness check to rule out the likelihood that mediated effects on academic test performance via language were simply due to the content overlap between the measures of early language performance and the reading subscales of the academic performance outcome measure at 7 years old. By way of comparison, the EF mediation models were also recomputed, using only the math subscale as the outcome, rather than the latent measure of academic test performance.

3.3 Results

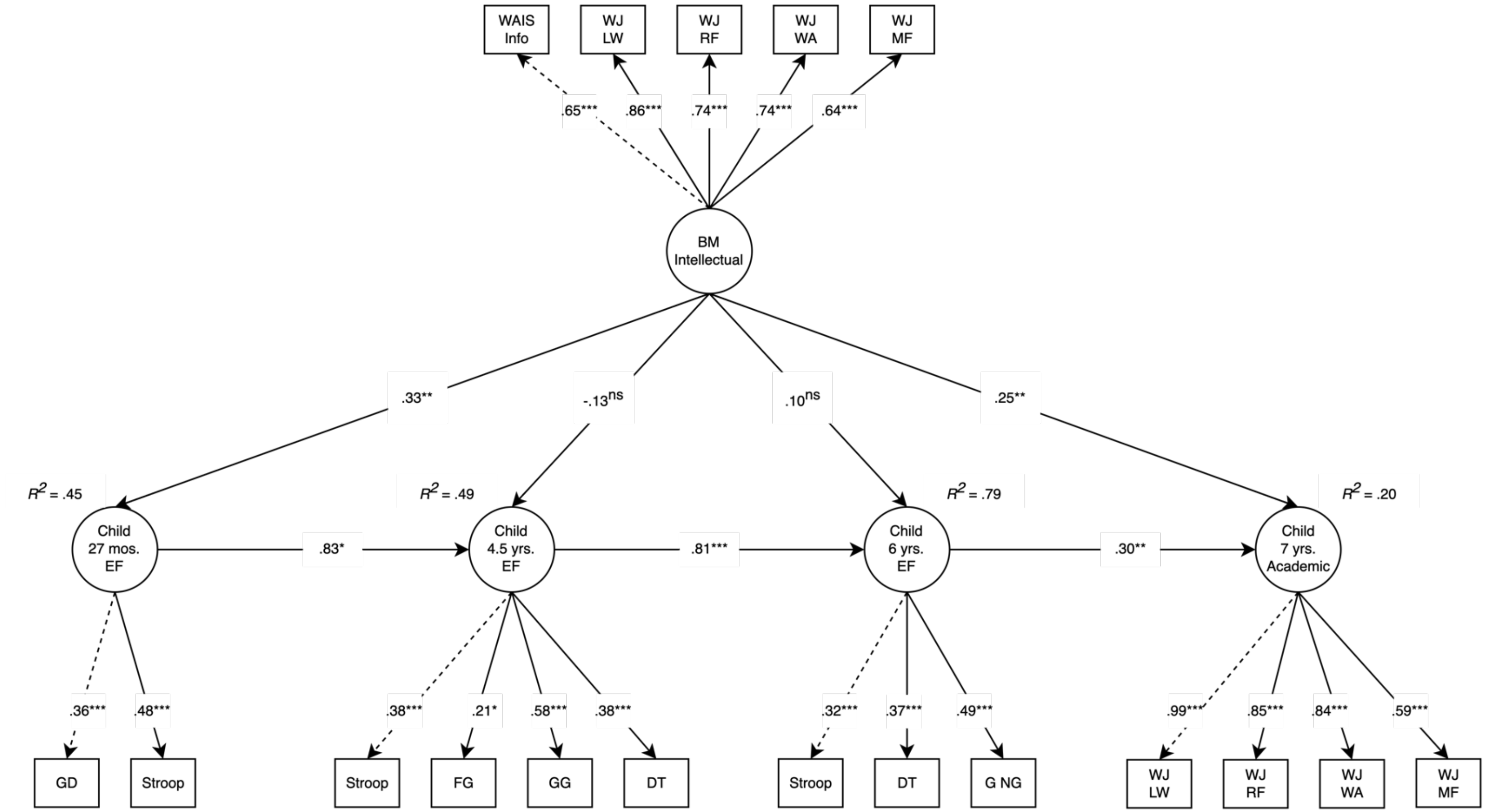
Means, standard deviations, sample sizes and bivariate correlations between study variables are presented in Table 3.1.

3.3.1 Early Executive Function and Later Academic Test Performance

Birth Mother Effects. As shown in Figure 3.2, birth mother intellectual performance was directly associated with EF at 27 months (β = .33, 95% CI [.11, .23], *p* = .004) and academic test performance at 7 years (β = .25, 95% CI [.10, .40], *p* = .001), and the total effect on academic test performance at 7 years was: β =.31, 95% CI [.19, .43], *p* < .001. There was no direct effect (β = −.13, 95% CI [−.51, .25], *p* = .507), indirect effect via EF at 27 months (β = .27, 95% CI [−.07, .61], *p* = .118), or total effect (β = .14, 95% CI [−.06, .34], *p* = .158) of birth mother intellectual performance on EF at 4.5 years. Nor was there a significant direct effect (β = .10, 95% CI [−.15, .34], *p* = .450) or total effect (β = .21, 95% CI [−.02, .44], *p* = .076) at 6 years. The model accounted for 45% of the variance in EF at 27 months, 49% of the variance in EF at 4.5 years, 79% of the variance in EF at 6 years and 20% of the variance in academic test performance at 7 years. The sensitivity analysis revealed that when the 27-month timepoint was dropped from the model, effects of birth mother intellectual performance did not carry forward to 4.5 years (Figure 1Ba, Appendix B). Nor did they carry forward to 6 years, when 27 months and 4.5 years were removed from the model (Figure 1Bb, Appendix B).

In the mediation model examining whether the effect of birth mother intellectual performance on academic test performance at 7 years was meditated via EF at 27 months, the total effect of birth mother intellectual performance on academic test performance at 7 years was statistically significant (β = .31, 95% CI [.19, .43], *p* < .001). The indirect effect, mediated through EF at 27 months, was 32% of the total effect and not statistically significant (β = .10, 95% CI [−.29, .49], *p* = .614). Model fit: $\chi^2(68) = 209.19$, *p* < .001, CFI = .91, RMSEA = .07, SRMR = .06.

Figure 3.2 Longitudinal Structural Equation Model Testing the Main Effects of Birth Mother Intellectual Performance on Child Executive Function and Academic Test Performance



Note. Model fit: $\chi^2(170) = 347.59$, $p < .001$, CFI = .90 RMSEA = .05, SRMR = .06. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother; EF = executive function; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; GD = gift delay; FG = forbidden gift; GG = guessing game; DT = dinky toys; G NG = Go NoGo. $^{ns}p \geq .1$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Effects on Math Performance. As in the original model that was being re-computed, in the sensitivity analysis re-computing the mediation analysis with the latent academic test performance variable at 7 years old replaced with the math fluency subscale of the Woodcock-Johnson, the indirect effect of birth mother intellectual performance on math test performance at 7 years old, mediated through child executive function at 27 months, was small and not statistically significant (β = .05, 95% CI [–.29, .39], *p* = .754). The indirect effect was 36% of the total effect and half the size (50%) of the indirect effect in the original model. Model fit: $\chi^2(52) = 160.94$, $p < .001$, CFI = .83, RMSEA = .09, SRMR = .06.

Birth Father Effects. As in the birth mother model, birth father intellectual performance directly predicted academic test performance at 7 years (β = .27, 95% CI [.05, .50], *p* = .017; Figure 3.3) and the total effect was: β = .34, 95% CI [.16, .52], *p* < .001. The effect estimate of birth father intellectual performance on EF at 27 months was numerically similar to the effect estimate in the birth mother model, although it was not statistically significant in the birth father model (β = .27, 95% CI [–.89, .63], *p* = .141). As in the birth mother model, birth father intellectual performance did not directly predict EF at 4.5 years (β = .16, 95% CI [–.27, .59], *p* = .474) or indirectly predict it via EF at 27 months (β = .20, 95% CI [–.13, .53], *p* = .232). However (unlike in the birth mother model), the total effect of birth father intellectual performance on EF at 4.5 years was significant (β = .36, 95% CI [.07, .65], *p* < .016). At 6 years, neither the direct effect (β = –.03, 95% CI [–.43, .37], *p* = .883) or the total effect (β = .26, 95% CI [–.04, .56], *p* = .094) reached statistical significance, and the effect estimates were similar to those in the birth mother model. The birth father model accounted for 43% of the variance in EF at 27 months, 55% of the variance in EF at 4.5 years, 73% of the variance in EF at 6 years and 19% of the variance in academic test performance at 7 years.

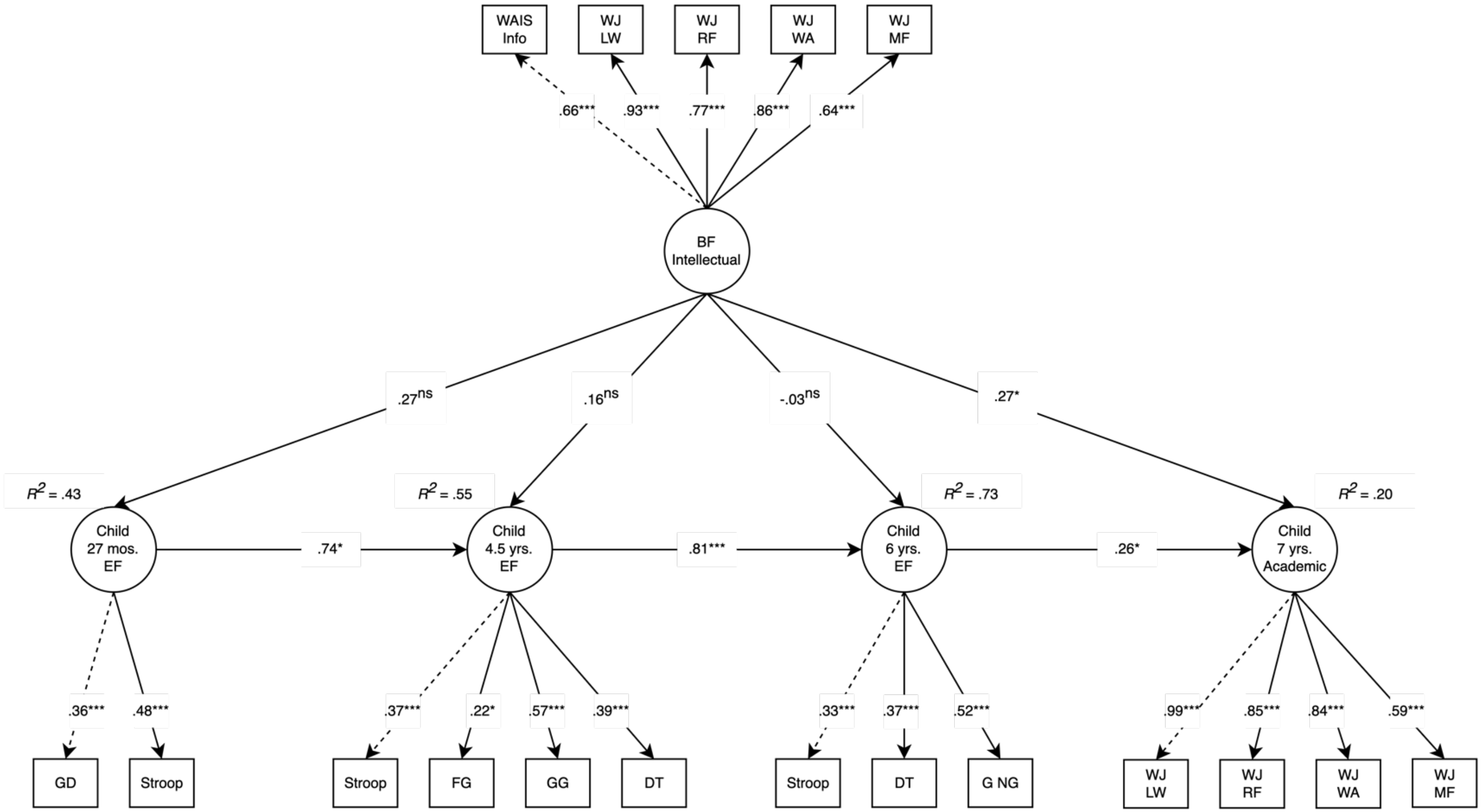
As in the birth mother mediation model, the total effect of birth father intellect on academic test performance at 7 years was statistically significant (β = .31, 95% CI [.13, .49], *p* = .001). The indirect effect of birth father intellectual performance on child academic test performance at 7 years, mediated through child EF at 27 months, was 22% of the total effect, not statistically significant (β = .07, 95% CI [–.22, .35], *p* = .640) and of comparable (small) magnitude to the birth mother results. Model fit: $\chi^2(68) = 120.12$, $p < .001$, CFI = .95, RMSEA = .04, SRMR = .07.

Effects on Math Performance. As in the original model that was being re-computed, in the sensitivity analysis re-computing the mediation analysis using the math fluency subscale of the Woodcock-Johnson at 7 years old (rather than the latent measure of academic test performance), the indirect effect of birth father intellectual performance on math test performance at 7 years old, mediated through child executive function at 27 months, was small and not statistically significant (β = .02, 95% CI [−.13, .18], $p = .768$). The indirect effect was 5% of the total effect and 29% the size of the indirect effect in the original model. Model fit: $\chi^2(36) = 49.98$, $p = .061$, CFI = .96, RMSEA = .03, SRMR = .07.

3.3.2 Early Language Performance and Later Academic Test Performance

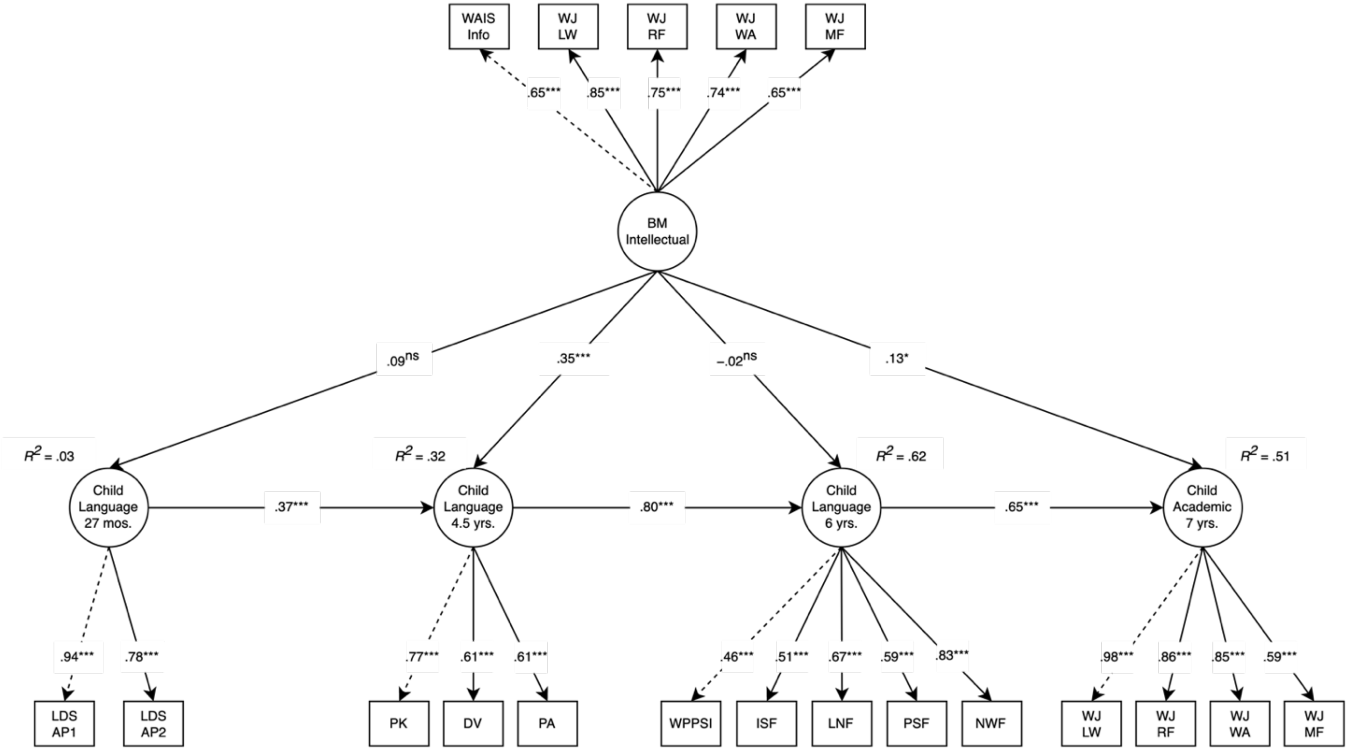
Birth Mother Effects. As displayed in Figure 3.4, birth mother intellectual performance directly predicted child language at 4.5 years (β = .35, 95% CI [.21, .49], $p < .001$) and the total effect at 4.5 years was: β = .38, 95% CI [.24, .53], $p < .001$. Birth mother intellectual performance also directly predicted child academic test performance at 7 years (β = .13, 95% CI [.02, .25], $p = .026$) and the total effect at 7 years was: β = .32, 95% CI [.20, .44], $p < .001$. There was no evidence of direct effects of birth mother intellectual performance on language at 27 months (β = .09, 95% CI [−.05, .22], $p = .201$) or 6 years (β = −.02, 95% CI [−.18, .13], $p = .768$), although there was a significant total effect at 6 years (β = .28, 95% CI [.14, .43], $p < .001$). The model accounted for 3% of the variance in language at 27 months, 32% of the variance in language at 4.5 years, 62% of the variance in language at 6 years and 51% of the variance in academic test performance at 7 years. A sensitivity analysis revealed that when language at 27 months and 4.5 years were dropped from the model, effects of birth mother intellectual performance carried forward to language at 6 years (Figure 2B, Appendix B).

Figure 3.3 Longitudinal Structural Equation Model Testing the Main Effects of Birth Father Intellectual Performance on Child Executive Function and Academic Test Performance



Note. Model fit: $\chi^2(170) = 347.59$, $p < .001$, CFI = .90 RMSEA = .05, SRMR = .06. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BF = birth father; EF = executive function; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; GD = gift delay; FG = forbidden gift; GG = guessing game; DT = dinky toys; G NG = Go NoGo. $^{ns}p \geq .1$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Figure 3.4 Longitudinal Structural Equation Model Testing Main Effects of Birth Mother Intellectual Performance on Child Language and Academic Test Performance



Note. Model fit: $\chi^2(190) = 403.06$, $p < .001$, CFI = .91, RMSEA = .05, SRMR = .07. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; LDS = Language Development Scale; AP1 = adoptive parent 1 report; AP2 = adoptive parent 2 report; PK = Test of Preschool Early Literacy (TOPEL) Print Knowledge; DV = TOPEL Definitional Vocabulary; PA = TOPEL Phonological Awareness; WPPSI = Wechsler Preschool and Primary Scale of Intelligence III; ISF = Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency; LNF = DIBELS Letter Naming Fluency; PSF = DIBELS Phoneme Segmentation Fluency; NWF = DIBELS Nonsense Word Fluency. $^{ns}p \geq .1$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

In the mediation model, examining whether the effect of birth mother intellectual performance on academic test performance at 7 years was meditated via language at 4.5 years, the total effect of birth mother intellectual performance on academic test performance at 7 years was statistically significant (β = .32, 95% CI [.20, .44], *p* < .001). The direct effect of birth mother intellectual performance on child academic test performance at 7 years was not statistically significant (β = .10, 95% CI [–.04, .24], *p* = .141) and the indirect effect, mediated through child language at 4.5 years, was statistically significant (β = .22, 95% CI [.08, .35], *p* = .002) and 68% of the total effect. Model fit: $\chi^2(81) = 211.74$, *p* < .001, CFI = .92, RMSEA = .06, SRMR = .06.

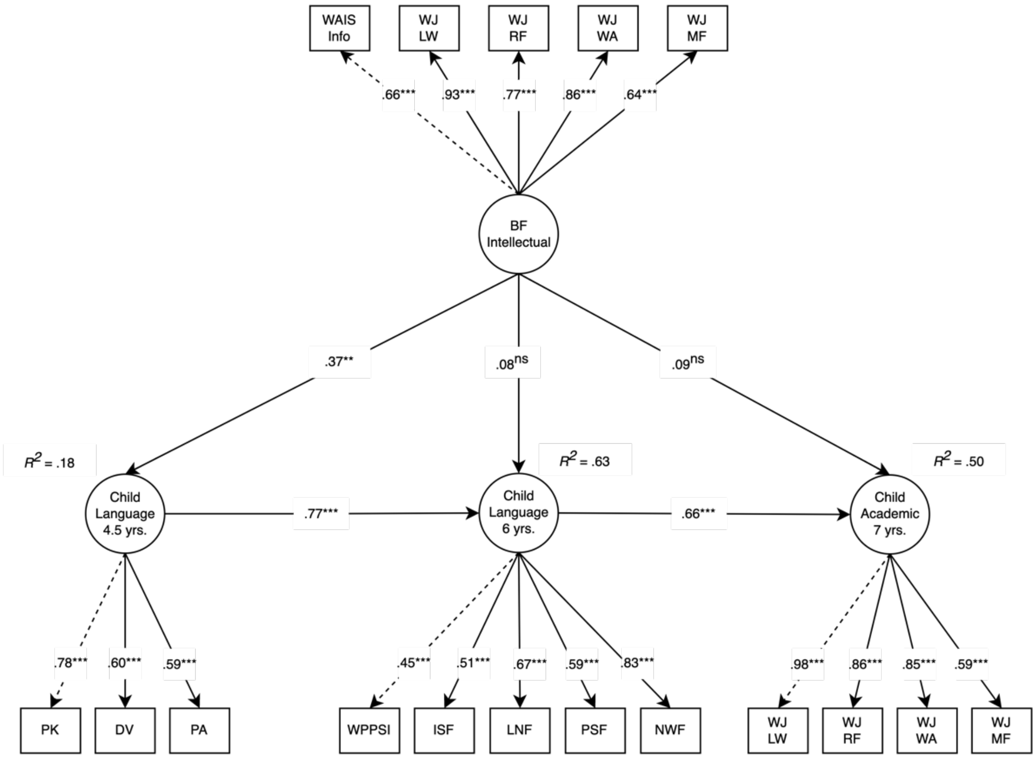
Effects on Math Performance. In the sensitivity analysis that re-computed the mediation analysis with the latent academic test performance variable at 7 years old replaced with the math fluency subscale of the Woodcock-Johnson, the findings were similar to those in the original model that was being re-computed. As in the original model, the indirect effect of birth mother intellectual performance on math test performance at 7 years old, mediated through child language at 4.5 years, was statistically significant (β = .14, 95% CI [.03, .24], *p* = .011). The indirect effect was 88% of the total effect and just under two thirds the size (64%) of the indirect effect in the original model. Model fit: $\chi^2(46) = 149.43$, *p* < .001, CFI = .87, RMSEA = .07, SRMR = .06.

Birth Father Effects. The model presented in Figure 3.4 was replicated in a sub-sample of children and their birth fathers. The model did not converge when the data at 27 months were included, so this timepoint was dropped from the model. As in the birth mother model, birth father intellectual performance significantly predicted child language at 4.5 years (β = .37, 95% CI [.11, .62], *p* = .005)—see Figure 3.5. Similar to the birth mother findings, there was no evidence of direct effects of birth father intellectual performance on language at 6 years (β = .08, 95% CI [–.21, .38], *p* = .575) and the total effect was significant (β = .36, 95% CI [.14, .60], *p* = .002). Unlike in the birth mother model, there was no evidence of direct effects of birth father intellectual performance on child academic test performance at age 7 years (β = .09, 95% CI [–.13, .30], *p* = .433), although, as in the birth mother model, the total effect at 7 years was significant (β = .33, 95% CI [.14, .51], *p* = .001). The model accounted for 18% of the variance in child language at 4.5 years, 63% of the variance in language at 6 years and 50% of the variance in academic test performance at 7 years. A sensitivity analysis revealed that, as in the birth mother sample, when language at 4.5 years was removed from the model, effects of birth father intellectual performance carried forward to language at 6 years (Figure 3Ba, Appendix B). When language at 4.5 and 6 years were dropped from the model, the effect of birth father intellectual performance on academic test performance at 7 years became significant (Figure 3Bb, Appendix B).

Similar to the birth mother mediation model, the total effect of birth father intellectual performance on academic test performance at 7 years was statistically significant ($\beta = .32$, 95% CI [.13, .50], $p = .001$). The direct effect of birth father intellectual performance on child academic test performance at 7 years old was not statistically significant ($\beta = .12$, 95% CI [−.09, .32], $p = .254$) and the indirect effect, mediated through child language at 4.5 years was statistically significant ($\beta = .20$, 95% CI [.04, .36], $p = .016$) and explained 63% of the total effect. The numerical estimates were similar to those in the birth mother model. Model fit: $\chi^2(81) = 132.20$, $p < .001$, CFI = .96, RMSEA = .04, SRMR = .07.

Effects on Math Performance. In the sensitivity analysis, re-computing the mediation model using only the math subscale at 7 years old, the effects of birth father intellectual performance continued to be mediated by language. As in the original model that was being re-computed, there was a significant indirect effect of birth father intellectual performance on child math performance at 7 years old, mediated via child language at 4.5 years ($\beta = .09$, 95% CI [.02, .16], $p = .018$). The indirect effect was 20% of the total effect and 45% the size of the indirect effect in the original model. Model fit: $\chi^2(46) = 54.55$, $p = .181$, CFI = .98, RMSEA = .02, SRMR = .07.

Figure 3.5 Longitudinal Structural Equation Model Testing Main Effects of Birth Father Intellectual Performance on Child Language and Academic Test Performance



Note. Model fit: $\chi^2(156) = 309.54$, $p < .001$, CFI = .91, RMSEA = .05, SRMR = .08. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BF = birth father; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = math fluency; PK = Test of Preschool Early Literacy (TOPEL) Print Knowledge; DV = TOPEL Definitional Vocabulary; PA = TOPEL Phonological Awareness; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; ISF = Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency; LNF = DIBELS Letter Naming Fluency; PSF = DIBELS Phoneme Segmentation Fluency; NWF = DIBELS Nonsense Word Fluency. $^{ns}p \geq .1$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

3.4 Discussion

Results were consistent with my hypothesis that effects of genetic influences on academic test performance at 7 years old are mediated by children’s early language. Birth mother and birth father general intellectual performance each predicted child language from 4.5 years onwards, but not at 27 months, and genetic effects on academic test performance at 7 years of age were mediated through language at 4.5 years. This is consistent with the large literature on genetic influences on children’s language (Stromswold, 2001) and extends the evidence by suggesting that language from 4.5 years old is an early manifestation of genetic influences on later intellectual performance. As the birth parent outcomes were measured in adulthood, the associations between birth mother or father intellectual performance and child language at 4.5 years are akin to “instant longitudinal” associations (Plomin, 1986), indicating that early language may be a marker of genetic effects, not only on academic test performance at 7 years, but also on general intellectual performance in adulthood.

The significant association between birth mother (but not birth father) intellectual performance and child executive function (EF) at 27 months, previously reported by Leve, DeGarmo, et al. (2013), was limited to this single occasion of measurement and did not reliably carry forward to 4.5 or 6 years. Although the high autocorrelations between the EF measures at the different timepoints suggest it is likely the effect at 27 months carried forward to later timepoints, the indirect and total effects of birth mother intellectual performance on child EF at 4.5 and 6 years were not statistically significant. In the birth father model, although none of the direct effects of birth father intellectual performance on child EF were significant, the total effect at 4.5 years old was statistically significant. Furthermore, as the birth father model was underpowered, I had anticipated paying greater attention to effect sizes than confidence intervals and *p* values. The effect sizes for the direct effect of birth father intellectual performance on child EF at 27 months and the total effects at 4.5 and 6 years were all moderately sized (β range: .27–.36), suggesting possible genetic effects on EF throughout early and middle childhood. The EF mediation models did not provide evidence of mediation of genetic effects on academic test performance at 7 years through EF at 27 months.

Evidence that language from 4.5 years old may be an early manifestation of genetic influences on later intellectual performance converges with findings from the polygenic score literature. For instance these findings are in line with results from the Born in Bradford study, reporting that genome-wide polygenic scores of total years of education achieved by adulthood (EA PGS) predicted a composite measure of academic test performance (including aspects of language) in 6–7-year-old school children (Armstrong-Carter et al., 2020). However, the present results provide evidence in a younger age group—preschool 4.5-year-old children. The absence of effects, in the EGDS sample, of birth parent intellectual performance on language at 27 months is at odds with detection in the Dunedin Study of a positive association between EA PGS and age of first words spoken, reported by parents when their children were 3 years old (Belsky et al., 2016). However, the present findings are consistent with evidence from the same study (Dunedin) that, while there was no association between children's EA PGS and their scores in the Peabody Picture Vocabulary Test at 3 years old, from 5 years onwards higher EA PGS predicted higher intelligence scores (captured by composite measures of verbal and non-verbal intelligence). The present findings are consistent with evidence that in infancy individual differences in language appear to be influenced to a greater degree by the shared environment than by genetic differences (Galsworthy et al., 2000b) but that by middle childhood, heritability of verbal and non-verbal cognitive performance is higher and the shared environmental component reduces (Davis et al., 2009a). The present results are also in line with evidence that the cross-time correlations for genetic influences on cognitive outcomes are low in early childhood and increase substantially across childhood (Tucker-Drob & Briley, 2014), as well as with evidence that from middle childhood the same genetic influences on cognitive skills predominate, increasing in magnitude as children get older (Briley & Tucker-Drob, 2013). As noted by Briley and Tucker-Drob (2013), one possible explanation for higher heritability of verbal and non-verbal cognitive performance by the time children reach school age is that when children enter formal schooling, standardized educational practices somewhat equalize environmental differences between them, allowing genetic differences to have greater influence on individual differences. An additional explanation—which is compatible with this study's findings, as well as with the reviewed literature on the increasing heritability of cognitive performance throughout childhood and increasing stability of genetic influences as children age—is that transactional mechanisms of gene-environment interplay amplify genetic effects through processes such as evocative and active gene-environment correlation (Scarr & McCartney, 1983).

3.4.1 Limitations and Future Directions

It is unclear whether the inconsistency of EF effects reflects a lack of effects of birth parent intellectual performance on child EF at later timepoints, and the absence of mediation of genetic effects on intellectual performance via EF, or a failure to operationalize EF sufficiently reliably at these occasions of measurement. Although the EF measures used in the present study were less internally consistent than the measures of language, the use of latent variables corrected for attenuation by error and the temporal stability of the EF latent variables was high. Compared to the temporal stability of language, the temporal stability of EF was higher from 27 months to 4.5 years and equivalent at 4.5 years to 6 years. It is also a possibility that EF was less predictive of later academic test performance than language due to high content overlap between indicators of language and the indicators of academic test performance that were included. However, this concern is somewhat mitigated by the results from the sensitivity analyses examining effects on only the math indicator of academic performance; the effects of birth parent intellectual performance continued to be mediated via language at 4.5 years old. This implies that language from 4.5 years is an early marker of genetic influences on a wider range of scholastic outcomes in middle childhood than simply those that are verbally oriented.

As my aim was to identify the earliest manifestations of genetic influences on later intellectual outcomes, it was important to include measures of EF and language from as early as 27 months in some of the analyses. However, as the 27-month measures miss important variance that is likely influenced by genetic pathways, estimates of effects on later child outcomes in the models that control for EF and language this early are substantially prone to omitted variable bias. Models not controlling for the earliest timepoint (which are thus less prone to this bias) are presented in Appendix B.

While these findings have the potential to aid the development of promotive and preventative interventions, they are unable to resolve uncertainty about whether early language is a liability-index (i.e. there are shared genetic factors that influence both language and subsequent academic test performance) or a causal mediator of genetic effects on subsequent academic test performance (i.e. limited verbal development would block the development of the skills necessary to perform well in academic tests) (Kendler & Neale, 2010). Each would have important but different implications for interventions in childhood. Although both suggest that low language is a risk factor for low academic test performance, the latter suggests that early intervention targeted at language might offset risk, whereas the former might be an indication in favour of more sustained support. Future research should be aimed at testing these alternatives, through longitudinal examination of academic test performance following interventions directly on early language.

It is a strength of the current analysis that it controlled for the influence of the prenatal environment, by including a measure of prenatal risk and through replicating the analyses in the birth father sample. However, the lack of statistical power to accurately estimate the influence of birth father genetic effects is a limitation. Sufficiently powered research is needed on the influence of birth father contributions to intellectual outcomes. Birth father models are not fully independent replications and almost all the measures of birth mother and birth father intelligence and academic test performance were correlated, suggesting the possibility of assortative mating, confounding, and partner interaction effects. In spite of the potential issues with spousal concordance, the birth father data add strength to the study—fathers play an equal role to mothers in contributing to the child's genotype, provide a control for intrapartum effects and are under-researched relative to mothers in developmental research. The role of birth fathers as a control for intrapartum effects is somewhat threatened by the potential for fathers to have indirect effects on foetal development through, for example, contributing to the family dynamics in the home, stress level of the mother and material resources accessible to the mother. However, the likelihood of this confounding the results is diminished by the fact that the rates of birth parent cohabitation in the sample were low.

All behaviour genetics findings represent 'what is' in a particular sample and cultural context rather than what 'could be' in a different context (Plomin et al., 2017). Consequently, it may be that there are features of the cultural milieu experienced by the US-based adopted children in the EGDS sample, that 'transmit' low-level genetic differences into differences in academic test performance to a greater or lesser degree than other cultural contexts might. Investigations into the representativeness of the EGDS sample have found that participating adoptive families appear to be representative of the US population (Leve, Neiderhiser, et al., 2013a). However, relative to the birth parents, they are higher socioeconomic status (SES) (Leve et al., 2019), which may bias findings. It cannot be assumed that the conclusions of this study hold for children reared in low SES environments, particularly as SES appears to moderate genetic effects on intellectual outcomes (Capron & Duyme, 1989; Tucker-Drob & Bates, 2015). There is evidence from the UK Biobank that EA PGS are more predictive of educational outcomes among non-adopted than adopted children, and that children in the lowest decile of polygenic score for educational attainment reach a significantly higher level of education if they are adopted than if they are not adopted (Cheesman et al., 2020). This converges with evidence from the US that children with low pre-adoption IQ scores experience substantial IQ score gains when adopted into high-SES families (Duyme et al., 1999), as well as with evidence that adoptees tend to academically out-perform their non-adopted biological siblings (Kendler et al., 2015). Collectively, these results indicate that genetic influences on education may be mediated by rearing environments or the wider cultural contexts that are associated with different rearing environments. Additionally, they suggest that estimates of direct genetic effects on academic outcomes may include mechanisms of gene-environment correlation and interaction, pointing to the possibility that genetic differences correlate and interact with different environmental mechanisms in different sociocultural contexts. There is evidence to suggest that different ethnic groups in the US and UK may exhibit different trajectories of verbal development (Saccuzzo et al., 1992; Zilanawala et al., 2016). For example, in the UK Millennium Cohort Study the ethnic groups in the sample had different odds of being in high or low performing profiles of verbal development in early childhood and these observed differences were mediated by the home learning environment, family routines, and the psychosocial environment (Zilanawala et al., 2016). Such findings illustrate the nuances of verbal development in different contexts and suggest that the present results might not hold in samples from different cultural and ethnic groups or socio-economic circumstances, within or outside of the US. It remains unclear how mechanisms of gene-environment interplay influence the development of academic outcomes in a diverse range of cultural contexts. Most behaviour genetics research—including the present study—is conducted in developed countries and majority White samples. Replication of these methods in other countries and sociodemographic groups is needed and until then it cannot be assumed that the present findings generalize to other cultural contexts. My interest in identifying a mediator in the association between birth parent and adopted offspring intellectual performance stems, in part, from an overarching aim to understand how rearing and learning environments may amplify the early manifestations of genetic influences on intellectual performance. However, it was not possible to form hypotheses about evocative effects of genetic influences underlying intellectual development without first identifying an early manifestation of genetic advantage that might elicit favourable and amplifying effects from parents. Now that language has been identified as a likely mediator of genetic influences on lifespan intellectual outcomes, we can posit early caregiving and learning conditions that might amplify genetic advantage. Children's language predicts parenting quality—including dimensions of parenting such as, sensitivity, positive regard, cognitive stimulation, and responsiveness—which in turn predicts reading performance (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012). Consequently, future research should explore whether these aspects of parenting amplify genetic advantage in language.

3.4.2 Conclusion

This is the first study (to my knowledge) to examine whether early EF or language mediate genetic effects on later intellectual performance. Effects of birth parent intellectual performance on child academic test performance at 7 years old were mediated through language at 4.5-years-old but were not mediated by early EF. These findings suggest that early language may be a manifestation of genetic advantage for lifespan intellectual outcomes. Based on the importance of intellectual performance for lifelong health and adjustment, the apparent role of early language in intellectual development represents a critical finding.

3.5 References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. University of Vermont, Research Center for Children, Youth & Families.

Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J. B., & Plomin, R. (2019). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry, 24*, 819–827. https://doi.org/10.1038/s41380-019-0394-4

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *The Journal of Abnormal Psychology, 112*, 545-557. https://doi.org/10.1037/0021-843X.112.4.545

Armstrong-Carter, E., Trejo, S., Hill, L. J. B., Crossley, K. L., Mason, D., & Domingue, B. W. (2020). The earliest origins of genetic nurture: The prenatal environment mediates the association between maternal genetics and child development. *Psychological Science, 31*(7), 781-791. https://doi.org/10.1177/0956797620917209

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., & Reiss, D. (2022). Early manifestations of intellectual performance: Evidence that genetic effects on later academic test performance are mediated through verbal performance in early childhood. *Child Development, 93*(2), e188-e206. https://doi.org/10.1111/cdev.13706

Austerberry, C., Mateen, M., Fearon, P., & Ronald, A. (2022). Heritability of psychological traits and developmental milestones in infancy: A systematic review and meta-analysis. *JAMA Network Open, 5*(8), e2227887-e2227887. https://doi.org/10.1001/jamanetworkopen.2022.27887

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

Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., Houts, R., Ramrakha, S., Sugden, K., Williams, B. S., Poulton, R., & Caspi, A. (2016). The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychololgical Science, 27*, 957-972. https://doi.org/10.1177/0956797616643070

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20*, 115-140. https://doi.org/10.2307/271084

Borriello, G. A., Ramos, A. M., Natsuaki, M. N., Reiss, D., Shaw, D. S., Leve, L. D., & Neiderhiser, J. M. (2020). The intergenerational transmission of mathematics achievement in middle childhood: A prospective adoption design. *Developmental Science, e12974*. https://doi.org/10.1111/desc.12974

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*(4498), 1055-1059. https://doi.org/10.1126/science.7195071

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24*(9), 1704-1713. https://doi.org/10.1177/0956797613478618

Capron, C., & Duyme, M. (1989). How much can we raise or lower I.Q.? The joint effects genetic and prenatal factors versus postnatal factors. *Journal de Genetic Human, 37*(2), 115-118.

Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., Eley, T. C., & Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological Science, 31*(5), 582-591. https://doi.org/10.1177/0956797620904450

Cioffi, C. C., Griffin, A. M., Natsuaki, M. N., Shaw, D. S., Reiss, D., Ganiban, J. M., Neiderhiser, J. M., & Leve, L. D. (2021). The role of negative emotionality in the development of child executive function and language abilities from toddlerhood to first grade: An adoption study. *Developmental Psychology, 57*, 347-360. https://doi.org/10.1037/dev0000972

Davis, O. S., Haworth, C. M., & Plomin, R. (2009a). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science, 20*(10), 1301-1308. https://doi.org/10.1111/j.1467-9280.2009.02433.x

Davis, O. S., Haworth, C. M., & Plomin, R. (2009b). Learning abilities and disabilities: generalist genes in early adolescence. *Cognitive Neuropsychiatry, 14*, 312-331. https://doi.org/10.1080/13546800902797106

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346. https://doi.org/10.1037/0033-295X.108.2.346

Domingue, B. W., & Fletcher, J. (2020). Separating measured genetic and environmental effects: Evidence linking parental genotype and adopted child outcomes. *Behavior Genetics, 50*, 301-309. https://doi.org/10.1007/s10519-020-10000-4

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428-1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of “dull children”?: A late adoption study. *Proceedings of the National Academy of Sciences, 96*(15), 8790-8794. https://doi.org/10.1073/pnas.96.15.8790

Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science, 3*, 206-215. https://doi.org/10.1111/1467-7687.00114

Ge, X., Natsuaki, M. N., Martin, D. M., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Villareal, G., Scaramella, L., Reid, J. B., & Reiss, D. (2008). Bridging the divide: Openness in adoption and postadoption psychosocial adjustment among birth and adoptive parents. *The Journal of Family Psychology, 22*, 529-540. https://doi.org/10.1037/a0012817

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*, 129-153. https://doi.org/10.1016/0010-0277(94)90068-x.

Goldsmith, H. H., & Rothbart, M. K. (1999). *The laboratory temperament assessment battery (LAB-TAB): Prelocomotor version, technical manual* (3.1 ed.). Department of Psychology, University of Wisconsin.

Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Institute for the Development of Educational Achievement.

Haworth, C. M., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J., van Beijsterveldt, C. E., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S., Kovas, Y., Corley, R. P., Defries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010, Nov). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*(11), 1112–1120. https://doi.org/https://doi.org/10.1038/mp.2009.55

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hummer, R. A., & Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States. *Population bulletin, 68*(1), 1-16.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227. https://doi.org/10.1080/02796015.1996.12085812

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. John Wiley & Sons.

Kendler, K. S., & Neale, M. C. (2010). Endophenotype: A conceptual analysis. *Molecular Psychiatry, 15*(8), 789-797. https://doi.org/10.1038/mp.2010.8

Kendler, K. S., Turkheimer, E., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proceedings of the National Academy of Sciences, 112*(15), 4612-4617. https://doi.org/10.1073/pnas.1417106112

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research, 41*, 627-641. https://doi.org/10.1044/jslhr.4103.627

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*, 490-507.

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*, 220-232.

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

Kosik, R., Mandell, G., Fan, A. P., Nguyen, T., Chen, J., & Eaton, W. (2018). The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study. *The European Journal of Psychiatry, 32*(2), 53-62. https://doi.org/10.1016/j.ejpsy.2018.01.001

Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*(10), 2048-2056. https://doi.org/10.1177/0956797613486982

Leve, L. D., DeGarmo, D. S., Bridgett, D. J., Neiderhiser, J. M., Shaw, D. S., Harold, G. T., Natsuaki, M. N., & Reiss, D. (2013). Using an adoption design to separate genetic, prenatal, and temperament influences on toddler executive function. *Developmental Psychology, 49*, 1045-1057. https://doi.org/10.1037/a0029390

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 22*(6), 716-727. https://doi.org/10.1017/thg.2019.66

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J., Natsuaki, M. N., & Reiss, D. (2013, Feb). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*(1), 412–423. https://doi.org/https://doi.org/10.1017/thg.2012.126

Lonigan, C. J., Allan, N. P., & Lerner, M. D. (2011). Assessment of preschool early literacy skills: Linking children's educational needs with empirically supported instructional activities. *Psychology in the Schools, 48*, 488-501. https://doi.org/10.1002/pits.20569

Lonigan, C. J., Wagner, R. K., & Torgesen, J. K. (2007). *Test of preschool early literacy examiner’s manual*. ProED.

Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children’s cognitive development across the first 3 years. *Child Development, 79*(4), 1065-1085. https://doi.org/10.1111/j.1467-8624.2008.01176.x

Marceau, K., De Araujo-Greecher, M., Miller, E. S., Massey, S. H., Mayes, L. C., Ganiban, J. M., Reiss, D., Shaw, D. S., Leve, L. D., & Neiderhiser, J. M. (2016). The perinatal risk index: Early risks experienced by domestic adoptees in the United States. *PLoS One, 11*, e0150486. https://doi.org/10.1371/journal.pone.0150486

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age

25 educational outcomes. *Early Childhood Research Quarterly, 28*, 314-324. https://doi.org/10.1016/j.ecresq.2012.07.008

McNeil, T. F., Cantor-Graae, E., & Sjostrom, K. (1994). Obstetric complications as antecedents of schizophrenia: Empirical effects of using different obstetric complication scales. *Journal of Psychiatric Research, 28*, 519-530. https://doi.org/10.1016/0022-3956(94)90042-6

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*, 308-330. https://doi.org/10.1016/j.dr.2010.07.001

Neiss, M., & Rowe, D. C. (2000). Parental education and child's verbal IQ in adoptive and biological families in the National Longitudinal Study of Adolescent Health. *Behavior Genetetics, 30*, 487-495. https://doi.org/10.1023/a:1010254918997

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-664. https://doi.org/10.1521/soco.19.6.625.20886

Plomin, R. (1986). Multivariate analysis and developmental behavioral genetics: Developmental change as well as continuity. *Behavior Genetics, 16*, 25-43. https://doi.org/10.1007/BF01065477

Plomin, R. (2018). *Blueprint: How DNA makes us who we are*. Allen Lane.

Plomin, R., Fulker, D. W., Corley, R. P., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 Years: A parent-offspring adoption study. *Psychological Science, 8*(6), 442-447. https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*, 592-617. https://doi.org/10.1037/0033-2909.131.4.592

Polderman, T. J. C., Posthuma, D., De Sonneville, L. M., Stins, J. F., Verhulst, F. C., & Boomsma, D. I. (2007). Genetic analyses of the stability of executive functioning during childhood. *Biological Psychology, 76*, 11-20. https://doi.org/10.1016/j.biopsycho.2007.05.002

Rescorla, L., & Alley, A. (2001). Validation of the language development survey (LDS): A parent report tool for identifying language delay in toddlers. *Journal of Speech, Language, and Hearing Research, 44*, 434-445. https://doi.org/10.1044/1092-4388(2001/035)

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/https://doi.org/10.18637/jss.v048.io2

Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal versus performance IQs for gifted African-American, Caucasian, Filipino, and Hispanic children. *Psychological Assessment, 4*(2), 239. https://doi.org/10.1037/1040-3590.4.2.239

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424-435. https://doi.org/10.2307/1129703

Selzam, S., Dale, P. S., Wagner, R. K., DeFries, J. C., Cederlof, M., O'Reilly, P. F., Krapohl, E., & Plomin, R. (2017). Genome-wide polygenic scores predict reading performance throughout the school years. *Scientific Studies of Reading, 21*, 334-349. https://doi.org/10.1080/10888438.2017.1299152

Sim, F., Thompson, L., Marryat, L., Ramparsad, N., & Wilson, P. (2019). Predictive validity of preschool screening tools for language and behavioural difficulties: A PRISMA systematic review. *PLoS One, 14*, e0211409. https://doi.org/10.1371/journal.pone.0211409

Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language, 77*, 647-723. https://doi.org/10.1353/lan.2001.0247

Tucker-Drob, E. M., & Bates, T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138-149. https://doi.org/10.1177/0956797615612727

Tucker-Drob, E. M., & Briley, D. A. (2014). Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychological Bulletin, 140*, 949-979. https://doi.org/10.1037/a0035893

Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene–environment transactions. *Developmental Science, 15*(2), 250-259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. Psychological Corporation.

Wechsler, D. (2002). *WPPSI-III administration and scoring manual*. Psychological Corporation.

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd, L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745-1761. https://doi.org/10.1111/cdev.13329

Wolchik, S. A., Tein, J. Y., Sandler, I. N., & Kim, H. J. (2016). Developmental cascade models of a parenting-focused program for divorced families on mental health problems and substance use in emerging adulthood. *Development and Psychopathology, 28*, 869-888. https://doi.org/10.1017/S0954579416000365

Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson tests of achievement (WJ III®)*. Riverside Publishing.

Yew, S. G., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*, 516-524. https://doi.org/10.1111/jcpp.12009

Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. E. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence, 69*, 8-15. https://doi.org/10.1016/j.intell.2018.03.011

Zilanawala, A., Kelly, Y., & Sacker, A. (2016). Ethnic differences in longitudinal latent verbal profiles in the millennium cohort study. *European Journal of Public Health, 26*(6), 1011-1016. https://doi.org/10.1093/eurpub/ckw184

Chapter 4: Evocative Effects on the Early Caregiving Environment of Genetic Influences Underlying the Development of Intellectual and Academic Abilities

4.1 Introduction

As outlined in **Chapter 3** (Austerberry, Fearon, et al., 2022) intellectual and academic performance are powerful predictors of health and success across the lifespan (Deary et al., 2010; Hummer & Hernandez, 2013; Kosik et al., 2018). They are also moderately to highly heritable, and their heritability appears to increase across the lifespan (Bouchard & McGue, 1981; C. M. A. Haworth et al., 2010; Kovas et al., 2013). Given the stability of the genome, this increasing heritability seems somewhat paradoxical. However, as discussed in **Chapter 1** of this thesis (Austerberry & Fearon, 2021), a plausible explanation for increasing heritability (and one that the present chapter aimed to examine in early intellectual and academic development) is that genetic differences are amplified across the lifespan as individuals influence, select and evoke differences in their environments that are correlated with their genotype (Plomin et al., 1977; Scarr & McCartney, 1983).

This thesis is particularly focused on early indicators of genetic and environmental influence. Given that young children have limited opportunities to select their environments, the form of gene-environment correlation (*r*GE) that may be particularly relevant during early childhood is evocative *r*GE, which occurs when an individual's genetically influenced characteristics systematically evoke differences in their environment (Plomin et al., 1977; Scarr & McCartney, 1983). In the context of intellectual and academic performance, children's academically relevant genetic differences may underlie early behavioural differences (e.g., vocabulary and interest in toys and books) that systematically elicit differences in the caregiving environment (e.g., warm, responsive parenting) that mediate genetic effects on academic outcomes, potentially amplifying initial genetic differences via environmental mechanisms. As these environmental mechanisms would be correlated with genetic differences, they could plausibly be masked by heritability estimates. This hypothesis is described in depth by Dickens and Flynn (2001) who explore the possibility that this amplification process accounts for rising levels of intelligence across the lifespan (and in successive cohorts of children and adults). They suggest that initially small genetic differences may become more potent over time through a multiplier effect produced by mechanisms of reciprocal causation between individuals' genotypes and the environmental influences they evoke. This explanation is challenging to test because it requires a genetically sensitive design and extensive longitudinal data. Consequently, to my knowledge, this amplification hypothesis has never been empirically tested in intellectual and academic development.

However, there is good reason to believe that children's early characteristics can elicit responses in their caregivers, evidenced by a robust body of literature that supports the existence of interplay between child traits and the parenting environment. This includes decades of evidence from phenotypic research of bidirectional effects between parents and children (Hipwell et al., 2008; Lugo-Gil & Tamis-LeMonda, 2008; Pardini et al., 2008) and the potential influence of child characteristics on parenting behaviour (Bell, 1968). Furthermore, behavioural genetics research provides evidence of possible evocative influences of children's genes on parenting. For example, there is a substantial literature from twin and parent-offspring adoption studies indicative of evocative effects of children's genes on differences in parenting (Klahr et al., 2017; Plomin & Bergeman, 1991), including in early childhood (Boivin et al., 2005; Elam et al., 2014; Fearon et al., 2015; Harold et al., 2013; Klahr et al., 2017; Knafo & Plomin, 2006a). However, only a small subset of this literature (outlined below) is specifically focused on intellectual and academic development.

Some phenotypic evidence suggests that children's intellectual differences evoke differences in parenting. For example, there is evidence of longitudinal bidirectional associations between children's cognitive ability and parental responsiveness, positive regard, cognitive stimulation, and sensitivity, between 14 months and 5 years of age (Blair et al., 2014; Lugo-Gil & Tamis-LeMonda, 2008). Furthermore, some genetically informative literature demonstrates that child → parent effects in intellectual development may be genetically driven. For example, Tucker-Drob and Harden (2012) found that, after controlling for parental cognitive stimulation at 2 years old, children's intellectual performance (a composite of verbal and nonverbal intelligence) at 2 years old predicted cognitively stimulating parenting when the children were 4 years of age. This association was almost entirely mediated by genetic variation, indicating that parents may adjust the way they parent in response to their children's genetic predispositions. Additionally, Wertz et al. (2020) found that children's polygenic scores for years of education (EduYears) predicted maternal positive parenting (cognitive stimulation, warmth and sensitivity) and home chaos when children were between 5 and 10 years of age, after controlling for mothers' own EduYears polygenic scores. However, evocative *r*GE in intellectual and academic development has never (to my knowledge) been examined using an adoption design. The present analyses aimed to address this by using a prospective parent-offspring adoption study to examine whether genetic influences that contribute to the development of children's intellectual ability evoke differences in the caregiving environment. The adoption design is ideally suited to investigate evocative *r*GE as, unlike all previous studies of evocative *r*GE in intellectual and academic development, it eliminates passive *r*GE,[1] while also testing the role of genetically mediated traits in evoking different responses within the family home.

I also aimed to test the Dickens and Flynn amplification hypothesis, by examining whether any observed differences in the caregiving environment would mediate genetic effects on later academic performance. To my knowledge, no previous research has tested whether observed evocative effects are responsible for increasing heritability of intellectual or academic outcomes. However, developmental research has uncovered several features of the rearing environment that predict intellectual and academic development. One of the most well-established constructs is positive parenting, which includes parental involvement, proactive anticipation of children's needs, responsivity, and warmth (Fan & Chen, 2001; Lugo-Gil & Tamis-LeMonda, 2008; Madigan et al., 2019). Other potentially important mechanisms include opportunities for learning and reading (Taylor et al., 2004), low levels of family chaos (Johnson et al., 2008; Petrill et al., 2004), and reduced use of screen media in the home environment (Adelantado-Renau et al., 2019), the latter including television, the internet, and video games. It is worth noting that the

---

[1] Passive *r*GE refers to the passive inheritance from biological parents of both genetic material and environments that are correlated with parental genotype, making it challenging to disentangle genetic from environmental effects in parent-child associations.

literature on screen time is somewhat mixed (Adelantado-Renau et al., 2019) compared to promotive effects of positive parenting. Based on this literature, I chose to focus my analyses primarily on positive parenting and secondarily (in exploratory analyses) on household chaos and screen time.

I also aimed to examine which genetically influenced child characteristics might evoke parenting differences. In **Chapter 3**, I established that children's language at 4.5 and 6 years of age appeared to be an early manifestation of genetic influences on intellectual ability at 7 years old and (as adoptee's language at 4.5 years of age was associated with their birth parent's intellectual ability in adulthood) potentially also into adulthood (Austerberry, Fearon, et al., 2022). Identifying early language as a likely mediator of genetic influences on lifespan intellectual outcomes informed my decision to examine whether children's language mediated any observed associations between genetic influences and parenting differences.

I tested the following three hypotheses. First, the hypothesis that genetic factors underlying academic abilities would have evocative effects on adoptive parent positive parenting in early childhood. Specifically, I expected that birth parent intellectual and academic performance (used as a proxy for genetic influences on children's intellectual and academic performance) would predict positive parenting when children were 4.5, 6, and 7 years old. Second, I tested the hypothesis that any observed evocative effects on adoptive parent parenting would predict children's subsequent language and academic test performance. Specifically, I expected adoptive parent parenting when children were 4.5 years and 6 years old would mediate the expected positive associations between birth parent intellectual performance and children's language at 6 and 7 years old. Third, I tested the hypothesis that language skills at 4.5 and 6 years of age would mediate the hypothesised associations between birth parent intellectual performance and adoptive parent positive parenting when children were 6 and 7 years of age.

4.2 Method

4.2.1 Participants

This study used data from the same pre-existing dataset as **Chapter 3**. Participants were 561 linked sets of adopted children and their birth mothers (*n* = 554), birth fathers (*n* = 210), and adoptive parents (562 adoptive fathers and 569 adoptive mothers) from the Early Growth and Development Study (EGDS), a US-based, longitudinal, prospective parent-child adoption study (Leve et al., 2019; Leve, Neiderhiser, et al., 2013a). The numbers of adoptive mothers and fathers do not sum to 561 because the sample includes 41 same-sex parent families and 15 additional adoptive parents who entered the family after the original couple adopted the child.

Participants were recruited through 45 adoption agencies, in 15 states across the US (Leve et al., 2019), in two cohorts: First, in 2003–2013, a sample of 361 adopted children and their birth and adoptive families and, second, in 2007–2017, 200 children and their families. EGDS assessments occurred in intervals of 9 months to 2 years and are still ongoing. I used data collected by EGDS from birth parents, at 18 months, 4.5 and 7 years postpartum, and adoptive parents and adoptees, when adoptees were 9, 18, and 27 months old, and 4.5, 6 and 7 years old. At each of these time points, two adoptive parents (referred to from now on as adoptive parent 1 and adoptive parent 2) were invited to participate.

The adopted children in the EGDS sample were 57% male. Adoptees were 56% non-Latinx White, 19% multiracial, 13% Black/African American, 11% Latinx and <1% “other” (including Asian, American Indian, and unknown ethnicity). The mean age at which the children were placed for adoption was 5.58 days postpartum (*SD* = 12.4; median = 2; range = 0–91). Adoptive parents were predominantly non-Latinx White (adoptive parent 1: 92%; adoptive parent 2: 90%). The remainder were Black/African American (adoptive parent 1: 4%; adoptive parent 2: 5%), Hispanic or Latinx (adoptive parent 1: 2%; adoptive parent 2: 2%), multiracial (adoptive parent 1: 1%; adoptive parent 2: 1%), and “other” (adoptive parent 1: 1%; adoptive parent 2: 2%). When the adoption took place, adoptive parents tended to be in their late-thirties (adoptive parent 1: *M* = 37.4, *SD* = 5.6; adoptive parent 2: *M* = 38.3, *SD* = 5.8), married or cohabiting (adoptive parent 1: 98%, adoptive parent 2: 100%), college educated (adoptive parent 1: 87%; adoptive parent 2: 81%) and with a combined median income above $100,000. Most of the birth parents were White (mothers: 70%; fathers: 70%). The remainder were Black/African American (mothers: 13%; fathers: 12%), Hispanic or Latinx (mothers: 7%; fathers: 10%), multiracial (mothers: 5%; fathers: 5%), and “other” (mothers: 5%; fathers: 4%). When the adoption took place, birth parents had a median household income of below $25,000, tended to be in their mid-twenties (mothers: *M* = 24.4, *SD* = 6.0; fathers: *M* = 26.1, *SD* = 7.8), not married or cohabiting (mothers: 6.1%; fathers: 14.0%), and not college educated (mothers: 75%; fathers: 84%). There is no evidence of selective placement in EGDS (Leve et al., 2019). Additional information about the recruitment, composition and representativeness of the study is reported elsewhere (Leve et al., 2019).

4.2.2 Ethics

Ethical approval was obtained from institutional review boards at the University of Oregon (Protocol number: 0304201400) and The Pennsylvania State University (Submission ID: CR00007591). Informed consent was obtained from all adult participants before research participation, and assent was obtained from children beginning at age 7 years.

4.2.3 Measures

Birth Parent General Intellectual Performance. As a proxy for genetic influences on adopted children’s intellectual development, I created the same five-indicator latent variable of birth parent general intellectual performance used in **Chapter 3** (Austerberry, Fearon, et al., 2022). The first item was a standardized total score on the 28-item Information subtest of the Wechsler Adult Intelligence Scale (Wechsler, 1997), administered to birth parents at 18 months postpartum. The remaining four indicators were T-scores from subtests of the Woodcock-Johnson Tests of Achievement III (Woodcock et al., 2001), administered to birth parents at 4.5 years postpartum: (1) 76-item Letter-Word Identification, measuring reading decoding; (2) 32-item Word Attack, capturing decoding and phonetic coding; (3) 98-item Reading Fluency, measuring reading speed and semantic processing speed; (4) 160-item Math Fluency, indexing math and numerical performance. Additional information on the internal consistency of these indicators and the rationale for combining them is reported in **Chapter 3** (Austerberry, Fearon, et al., 2022).

Child Language. I created the same latent variables measuring language at 4.5 years and 6 years as **Chapter 3** (Austerberry, Fearon, et al., 2022). The latent variable at 4.5 years had three indicators, each of which was a subscale from the Test of Preschool Early Literacy (TOPEL) (Lonigan et al., 2011): (1) 36-item Print Knowledge, measuring knowledge of the alphabet, written language conventions and written form; (2) 35-item Definitional Vocabulary, assessing definitional and single-word oral vocabulary; (3) 27-item Phonological Awareness, measuring word elision and blending. The latent variable at 6 years had five indicators, one of which was a standardized score from the vocabulary assessment of the Wechsler Preschool and Primary Scale of Intelligence III (Wechsler, 2002), and four of which were assessments from the Dynamic Indicators of Basic Early Literacy Skills (Good & Kaminski, 2002): (1) 16-item Initial Sound Fluency (ISF), measuring phonemic awareness; (2) Letter Naming Fluency (LNF), capturing proficiency in naming upper- and lower-case letters, using a list of 110 letters; (3) Phoneme Segmentation Fluency (PSF), assessing proficiency in fluently segmenting three- and four-phoneme words into their individual phonemes, using a list of 24 words; (4) Nonsense Word Fluency (NWF), testing understanding of the alphabetic principle, including letter-sound correspondence, using a list of 50 nonsense words. The psychometric justification for using these measures at the ages analysed is provided in **Chapter 3** of this thesis (Austerberry, Fearon, et al., 2022) and the references cited for each measure. Further information on the longitudinal stability, reliability and internal consistency of the language measures and indicators is also reported in **Chapter 3** (Austerberry, Fearon, et al., 2022).

Child Academic Test Performance. I created the same latent variable to estimate child academic test performance at 7 years old as in **Chapter 3** (Austerberry, Fearon, et al., 2022), drawing on the same four indicators of reading and math performance that were administered to birth parents from the Woodcock-Johnson Tests of Achievement III (Woodcock et al., 2001).

Caregiving Environment. Prior to hypothesis testing, I conducted a split-half exploratory factor analysis (EFA) of items from the following measures administered to adoptive parents when the children were 4.5, 6 and 7 years old: (1) the Chaos, Hubbub, and Order Scale (CHAOS) (Matheny et al., 1995), (2) the Home Literacy Environment (HLE) questionnaire (Niklas & Schneider, 2013), (3) the Involvement and Positive Parenting subscales of the Alabama Parenting Questionnaire (APQ) (Shelton et al., 1996a), and (4) the Warmth subscale of the Iowa Family Interaction Rating Scales (IOWA) (Melby & Conger, 2001). The results from the EFA are reported in Appendix C. Based on the factor structure of the items, I used the following three latent variables in the analyses (the first of which was used in the main analysis, and the latter two in exploratory analyses):

Parental Warmth. I created a latent variable measuring positive parenting at 4.5, 6, and 7 years using the six self-report indicators from the Warmth subscale of the IOWA (Melby & Conger, 2001). Each item was scored on a seven-point scale from 1 (Never) to 7 (Always). The psychometric information for the IOWA is reported by Melby and Conger (2001). The internal consistency of the items in the EGDS sample was good (α = .87, .86, .88 for adoptive parent 1 ratings at 4.5, 6, and 7 years, respectively; α = .85, .87, .87 for adoptive parent 2 ratings).

Home Chaos. I constructed a latent variable measuring household chaos at ages 4.5, 6 and 7 years using three indicators from adoptive parent 1 CHAOS ratings (Matheny et al., 1995): (1) “You can’t hear yourself think in our home”, (2) “It’s a real zoo in our home”, (3) “The atmosphere in our house is calm” (reverse scored). Each item was rated on a 5-point scale from 1 (Definitely Untrue) to 5 (definitely true). The psychometric information for this measure is provided by Matheny et al. (1995). The internal consistency of the three items was good in the EGDS sample at each time point (α = .81, .82, .85 at 4.5, 6 and 7 years, respectively).

Screen Media Use**.** I created a latent variable measuring screen media use in the household at ages 4.5, 6, and 7 years using four indicators from adoptive parent 1 ratings. Three items were from the HLE questionnaire (Niklas & Schneider, 2013): “On average, how many hours per day does your child watch television or play video games?” on (1) “Weekdays,” (2) “Saturday,” (3) “Sunday”). One indicator was from the CHAOS: “There is usually a television turned on somewhere in our home” (rated on a 5-point scale from 1 [Definitely Untrue] to 5 [definitely true]). The psychometric information for the HLE is reported by Niklas and Schneider (2013) and the psychometric information for the CHAOS is reported by (Matheny et al., 1995). The internal consistency of the four items in the EGDS sample was good (α = .83, .79, .82 at 4.5, 6 and 7 years, respectively).

Covariates. The same covariates were used in the present study as **Chapter 3** (Austerberry, Fearon, et al., 2022): adoption openness, child sex, and prenatal risk. Adoption openness was measured by EGDS using a 4-item measure (Ge et al., 2008), averaged across ratings provided at 9, 18, and 27 months postpartum by birth mothers and adoptive parents. I used a weighted prenatal risk score created by EGDS based on the work by Marceau et al. (2016) drawing on birth mother reports of maternal and pregnancy complications, labour and delivery complications and neonatal complications at 5 months postpartum (McNeil et al., 1994).

4.2.4 Data Analysis

The analyses were preregistered with the Open Science Framework in March 2021 (Austerberry et al., 2021). I conducted longitudinal cross-lagged panel modelling (CLPM) in the lavaan package (Rosseel, 2012) in R 4.1.2, using structural equation modelling, which combines a measurement model (also known as confirmatory factor analysis [CFA]) with a structural model testing the proposed causal relations. The full CLPM was built in several steps: First, I constructed an autoregressive model examining parental warmth at 4.5, 6, and 7 years old, in which warmth at 7 years was regressed on warmth at 6 years, which was regressed on warmth at 4.5 years. Second, I introduced birth parent intellectual ability (a proxy for genetic influence) as a predictor in the autoregressive models, examining whether the birth parent general intellectual ability latent variable predicted adoptive parent warmth at 4.5, 6, and 7 years. Third, I constructed a CLPM, combining the parenting autoregressive models and genetic predictor with the autoregressive model examining children's language and academic performance reported in **Chapter 3** (Austerberry, Fearon, et al., 2022). The following cross-lagged associations between parental warmth and children's language and academic performance were included in the model: child academic test performance at 7 years old was regressed on parental warmth at 6 years old; child language at 6 years old was regressed on warmth at 4.5 years old; parental warmth at 7 years old was regressed on language at 6 years old; and warmth at 6 years old was regressed on language at 4.5 years old. Additionally, language and warmth were allowed to covary at each timepoint. I preregistered a fourth step in the analysis—to introduce random intercepts to the CLPMs, in line with (Hamaker et al., 2015), separating within-person variation from between-person differences. When I introduced random intercepts to our models, none of them converged. These convergence issues may have been evident because it is not appropriate in a model that already has a built-in control for the presence of time-invariant covariates (the random intercept) to introduce time-invariant variables (e.g., birth parent intellectual ability, which is a proxy for genetic influences and remains constant over time) as predictors of within-person variation.

The main analysis was on parenting warmth based on an established evidence base (reviewed in the Introduction to this chapter) to support the hypothesis that positive parenting may be implicated in evocative processes in educational development. Exploratory analyses were also conducted on home chaos and screen media use. Separate models were run on parental warmth data from adoptive parent 1 and adoptive parent 2. Primary analyses were conducted using data from birth mothers because it is a larger sample than the sample of birth fathers. I conducted semi-independent replications of these models using a smaller sample of birth fathers, providing an estimate of genetic effects that is not confounded by prenatal effects. While the birth father sample is the largest ever recruited in a prospective parent-offspring adoption study, it has reduced statistical power compared to birth mother analyses. Although determining sample size requirements for complex structural equation models is not straightforward, generally a minimum sample size of 200 (Kline, 2016) or a sample size of five to ten times the number of observed variables is considered to be acceptable (Bentler & Chou, 1987; Nunnally, 1967). As the main analyses contained 29 observed variables, this would suggest a sample size of at least 145-290. Thus, I anticipated that comparisons between results from birth mothers ($n = 325$) and birth fathers ($n = 109$) would focus on the magnitude of the path coefficients rather than on p values or confidence intervals. As in **Chapter 3** (Austerberry, Fearon, et al., 2022), based on recommendations by Hu and Bentler (1999), we used a combination rule, according to which model fit was considered adequate if SRMR < .09 and RMSEA < .06. As in **Chapter 3** (Austerberry, Fearon, et al., 2022), the indirect effects were estimated using bootstrapping with 5000 repetitions (Bollen & Stine, 1990).

4.2.5 Missing Data

Sample sizes for the variables used in the main analyses are reported in Table 4.1 and for those in the exploratory analyses reported in Table 4.2. The primary source of missing data in models using birth mother data was children's language scores at 6 years old. In birth father models, the primary source of missing data was missing information on birth father intellectual performance. The data were not missing completely at random (MCAR [Little's MCAR χ^2 (5635) = 6857, $p < .01$]). MCAR occurs when the probability of being missing is the same for all cases, and there is no systematic association between the missingness of the data and any other (observed or missing) values. As in **Chapter 3** (Austerberry, Fearon, et al., 2022), I ran an attrition analysis using the Missing Value Analysis (MVA) function in SPSS, which uses the t-test procedure to compare group means and patterns of missingness in the data. This analysis revealed that the patterns of missingness for almost all (97%) of the variables used in the analyses were related to the observed values of other variables in the dataset. It was not possible to completely rule out the possibility that the data were missing not at random (MNAR) because that would require measuring the missing data (e.g., through following up non-respondents). However, the observed patterns of missingness are consistent with the data being missing at random (MAR), which occurs when the missingness of a variable is systematically related to the observed but not unobserved data. As in **Chapter 3** (Austerberry, Fearon, et al., 2022), missing data were handled using full information maximum likelihood (FIML), which is suitable for data that are MAR and is of comparable performance to multiple imputation (Allison, 2003).

Table 4.1 Means, Standard Deviations and Sample Sizes of Study Variables

	Birth Parent General Intellectual Performance					
	Birth Mother			Birth Father		
Variable	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
WAIS Information	323	9.56	2.59	102	10.65	2.88
WJ Letter-Word Frequency	325	47.89	5.85	109	47.29	7.67
WJ Word Attack	325	49.80	7.53	109	46.6	7.80
WJ Reading Fluency	325	46.87	6.73	109	46.94	8.40
WJ Math Fluency	325	44.04	8.86	109	41.95	10.01
	Adoptive Parent Warmth					
	Adoptive Parent 1			Adoptive Parent 2		
Variable	*n*	*Mean*	*SD*	*n*	*Mean*	*SD*
4.5 yrs. Let him/her know you really care about him/her	414	6.57	0.65	374	6.41	0.70
4.5 yrs. Act loving and affectionate toward him/her	414	6.55	0.71	374	6.47	0.61
4.5 yrs. Let your child know that you appreciate him/her, his/her ideas, or things he/she does	414	6.40	0.83	374	6.25	0.83
4.5 yrs. Help him/her do something that was important to him/her	414	6.14	0.98	374	6.03	0.85
4.5 yrs. Act supportive and understanding toward him/her	414	6.32	0.83	374	6.31	0.75
4.5 yrs. Tell him/her you love him/her	399	6.79	0.56	361	6.71	0.61
6 yrs. Let him/her know you really care about him/her	403	6.50	0.81	362	6.35	0.78
6 yrs. Act loving and affectionate toward him/her	403	6.54	0.69	362	6.38	0.76
6 yrs. Let your child know that you appreciate him/her, his/her ideas, or things he/she does	403	6.41	0.77	362	6.19	0.85
6 yrs. Help him/her do something that was important to him/her	403	6.17	0.82	362	5.97	0.92
6 yrs. Act supportive and understanding toward him/her	403	6.38	0.69	362	6.26	0.71
6 yrs. Tell him/her you love him/her	403	6.78	0.57	362	6.63	0.68
7 yrs. Let him/her know you really care about him/her	307	6.48	0.71	273	6.24	0.86
7 yrs. Act loving and affectionate toward him/her	307	6.46	0.68	273	6.29	0.78
7 yrs. Let your child know that you appreciate him/her, his/her ideas, or things he/she does	306	6.30	0.80	273	6.09	0.94
7 yrs. Help him/her do something that was important to him/her	307	6.02	0.88	273	5.79	0.92
7 yrs. Act supportive and understanding toward him/her	307	6.33	0.71	273	6.12	0.83
7 yrs. Tell him/her you love him/her	307	6.75	0.54	273	6.56	0.76
	Child language and academic performance					
Variable	*n*	*Mean*	*SD*			
4.5 yrs. TOPEL Print Knowledge	291	106.80	13.39			
4.5 yrs. TOPEL Definitional Vocabulary	288	104.30	9.94			
4.5 yrs. TOPEL Phonological Awareness	137	97.76	15.66			
6 yrs. WPPSI	288	10.31	2.28			
6 yrs. DIBELS Initial Sound Fluency	141	58.48	30.22			
6 yrs. DIBELS Letter Naming Fluency	288	52.18	31.68			
6 yrs. DIBELS Phoneme Segmentation Fluency	293	33.84	27.87			
6 yrs. DIBELS Nonsense Word Fluency	293	46.19	32.58			
7 yrs. WJ Letter-Word Frequency	336	56.29	9.23			
7 yrs. WJ Word Attack	334	53.91	10.80			
7 yrs. WJ Reading Fluency	334	55.22	7.51			
7 yrs. WJ Math Fluency	334	49.66	10.15			

Note. SD = Standard deviation. WJ = Woodcock-Johnson Tests of Achievement III. WAIS = Wechsler Adult Intelligence Scale-III. TOPEL = Test of Preschool Early Literacy. WPPSI = Wechsler Preschool and Primary Scale of Intelligence III. DIBELS = Dynamic Indicators of Basic Early Literacy Skills.

Table 4.2 Means, Standard Deviations and Sample Sizes of Variables in Exploratory Analyses

	Adoptive Parent 1 Rated Chaos		
Variable	n	Mean	SD
4.5 yrs. You can’t hear yourself think in our home	417	2.25	1.14
4.5 yrs. It’s a real zoo in our home	417	2.07	1.09
4.5 yrs. The atmosphere in our house is calm (reverse scored)	417	2.42	1.06
6 yrs. You can’t hear yourself think in our home	391	2.20	1.13
6 yrs. It’s a real zoo in our home	391	2.07	1.08
6 yrs. The atmosphere in our house is calm (reverse scored)	391	2.37	1.02
6 yrs. You can’t hear yourself think in our home	396	2.25	1.13
6 yrs. It’s a real zoo in our home	396	2.14	1.12
6 yrs. The atmosphere in our house is calm (reverse scored)	396	2.43	0.98

	Adoptive Parent 1 Rated Screen Use		
Variable	n	Mean	SD
4.5 yrs. On average, how many hours per day does your child watch television or play video games on weekdays?	415	1.64	1.23
4.5 yrs. On average, how many hours per day does your child watch television or play video games on Saturday?	415	2.09	1.27
4.5 yrs. On average, how many hours per day does your child watch television or play video games on Sunday?	415	1.90	1.25
4.5 yrs. There is usually a television turned on somewhere in our home	417	2.55	1.40
6 yrs. On average, how many hours per day does your child watch television or play video games on weekdays?	252	1.41	1.02
6 yrs. On average, how many hours per day does your child watch television or play video games on Saturday?	252	2.13	1.17
6 yrs. On average, how many hours per day does your child watch television or play video games on Sunday?	252	1.92	1.24
6 yrs. There is usually a television turned on somewhere in our home	391	2.46	1.36
7 yrs. On average, how many hours per day does your child watch television or play video games on weekdays?	298	1.42	1.25
7 yrs. On average, how many hours per day does your child watch television or play video games on Saturday?	298	2.40	1.29
7 yrs. On average, how many hours per day does your child watch television or play video games on Sunday?	298	2.28	1.33
7 yrs. There is usually a television turned on somewhere in our home	396	2.51	1.38

Note. SD = Standard deviation.

4.3 Results

Descriptive statistics (sample sizes, means and standard deviations) for the variables used in the main analyses are presented in Table 4.1 and descriptive statistics for the exploratory analyses are in Table 4.2.

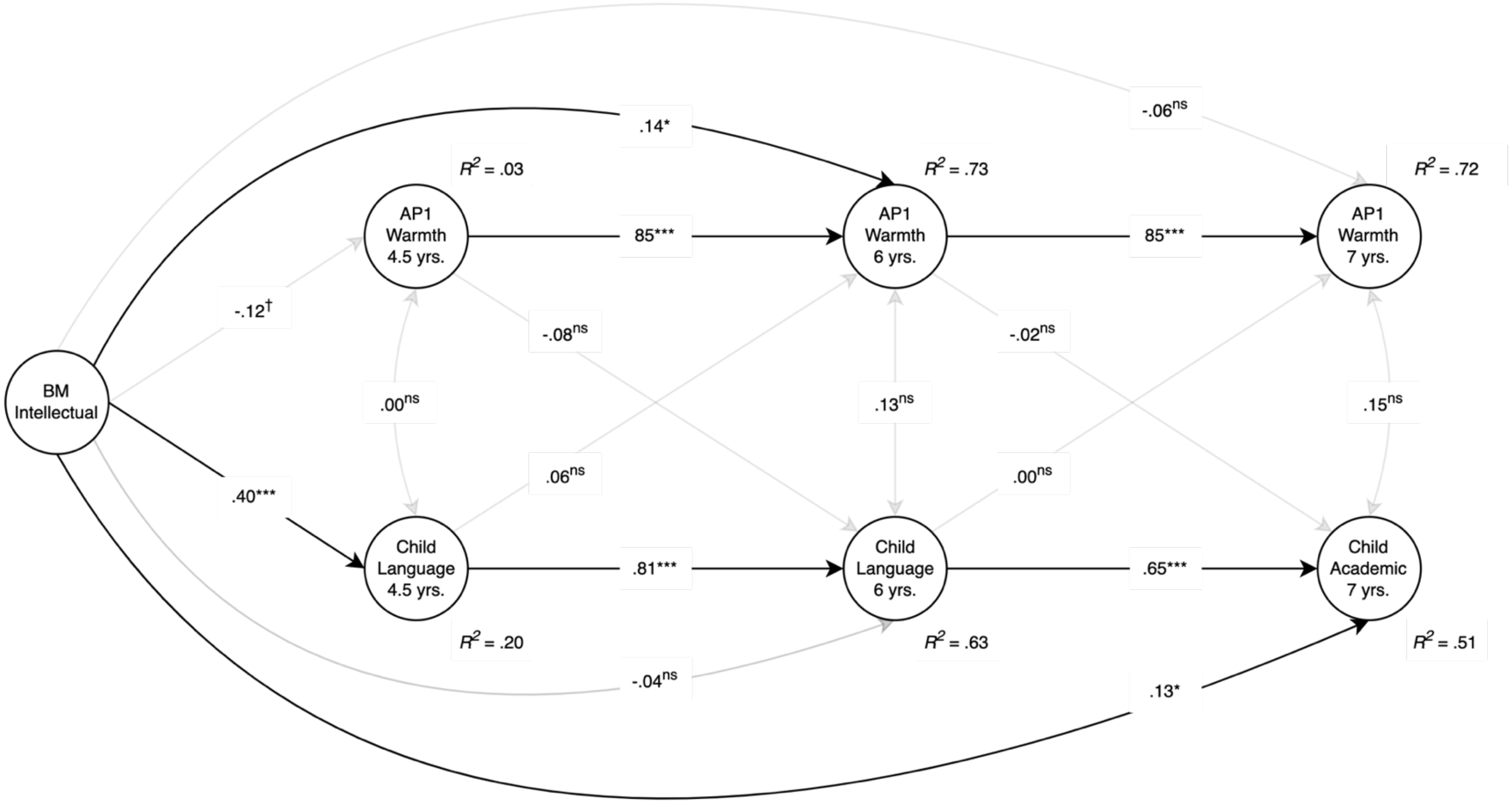
4.3.1 Birth Mother Effects on, and Cross-lagged Associations Between, Adoptive Parent Warmth and Children's Language and Academic Performance

Adoptive Parent 1.

Birth Mother Effects on Warmth. As displayed in Figure 4.1, there was a significant direct effect of birth mother intellectual performance (a proxy for genetic influences) on adoptive parent 1 warmth at 6 years old (β = .14, 95% CI [.01, .27], p = .038). The direct effect of birth mother intellectual performance on adoptive parent 1 warmth was not statistically significant at 4.5 years (β = –.12, 95% CI [–.26, .01], p = .063) or 7 years (β = –.06, 95% CI [–.17, .06], p = .361). However, the indirect effect of birth mother intellectual performance on adoptive parent 1 warmth at 7 years via warmth at 6 years was statistically significant (β = .12, 95% CI [.01, .23], p = .040). The model accounted for 3% of the variance in adoptive parent 1 warmth at 4.5 years, 73% of the variance in warmth at 6 years and 72% of the variance in warmth at 7 years.

Cross-lagged Associations. As shown in Figure 4.1, the direct effect of adoptive parent 1 warmth at 4.5 years on child language at 6 years old was not statistically significant (β = –.08, 95% CI [–.21, .05], p = .208) and nor were the effects of warmth at 6 years on child academic test performance at 7 years (β = –.02, 95% CI [–.12, .09], p = .780), child language at 4.5 years on warmth at 6 years (β = .06, 95% CI [–.08, .20], p = .413), or child language at 6 years on warmth at 7 years (β = .00, 95% CI [–.11, .12], p = .027). The indirect effect of birth mother intellectual performance on adoptive parent 1 warmth at 6 years via child language at 4.5 years was not statistically significant (β = .02, 95% CI [–.03, .08], p = .418). Nor was there evidence of indirect effects of birth mother intellectual performance on adoptive parent 1 warmth at 7 years old via earlier language, or of indirect effects of birth mother intellectual performance on child language at 6 years or academic performance at 7 years old via earlier parental warmth.

Figure 4.1 Longitudinal Structural Equation Model Examining the Effects of Birth Mother Intellectual Performance on, and Cross-lagged Associations Between, Adoptive Parent 1 Warmth and Children's Language and Academic Performance



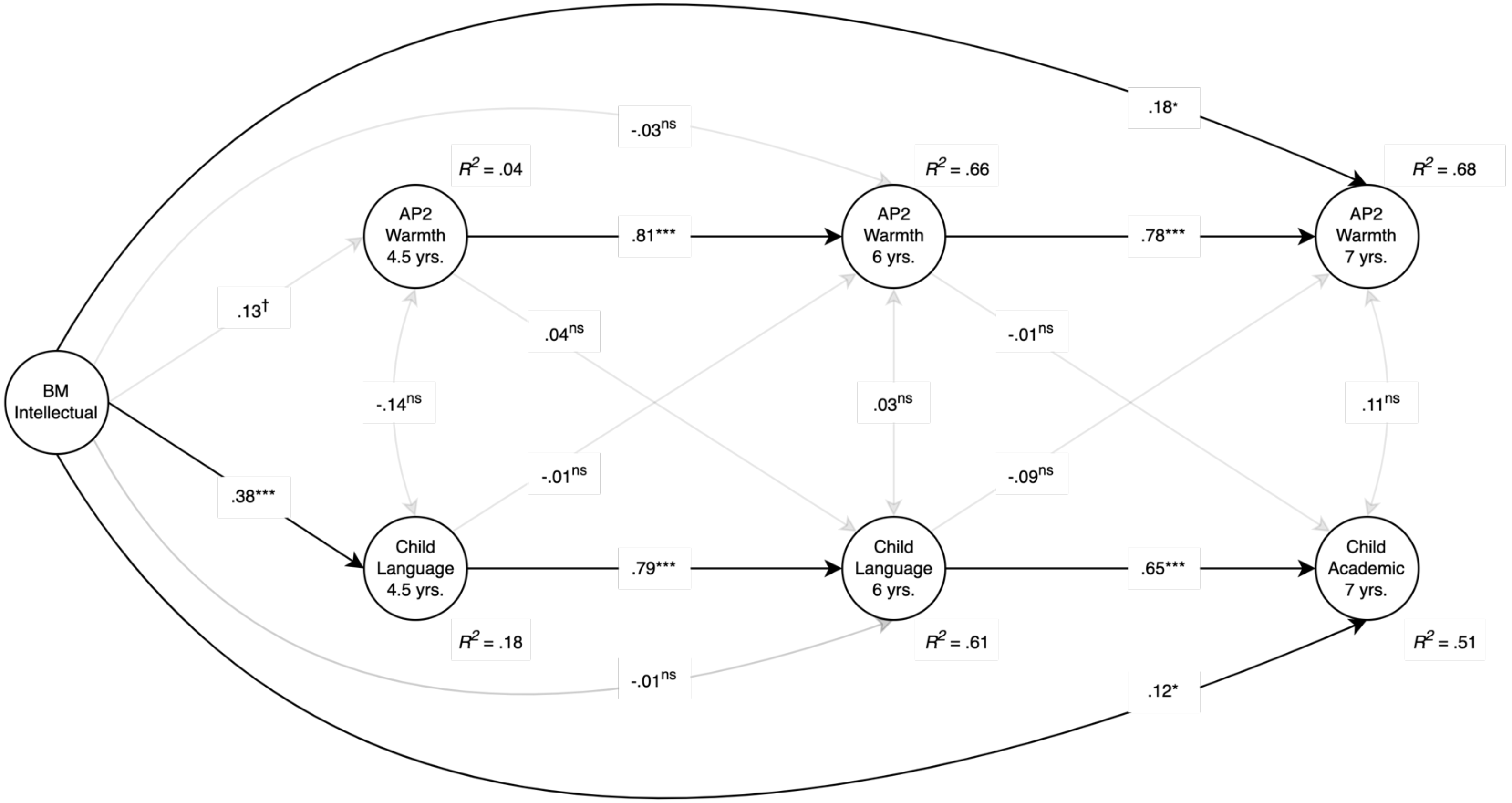
Note. Model fit: $\chi^2(630) = 1346$, $p < .001$, CFI = .88, RMSEA = .05, SRMR = .07. Standardized estimates reported. Solid paths represent significant associations ($p < .05$). Faded paths represent non-significant associations ($p \geq .05$). Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother. AP1 = adoptive parent 1. $^{ns}p \geq .1$. $^{†}\ p < .1$. $^{*}p < .05$. $^{***}p < .001$.

Adoptive Parent 2

Birth Mother Effects on Warmth. As displayed in Figure 4.2, there was a significant direct effect of birth mother intellectual performance on adoptive parent 2 warmth at 7 years old (β = .18, 95% CI [.05, .31], p = .007). The direct effect of birth mother intellectual performance on adoptive parent 2 warmth was not statistically significant at 4.5 years (β = .13, 95% CI [−.02, .28], p = .087) or at 6 years (β = −.03, 95% CI [−.18, .11], p = .661).The model accounted for 4% of the variance in adoptive parent 2 warmth at 4.5 years, 66% of the variance in warmth at 6 years and 68% of the variance in warmth at 7 years.

Cross-lagged Associations. As displayed in Figure 4.2, the direct effect of adoptive parent 2 warmth at 4.5 years on child language at 6 years old was not statistically significant (β = .04, 95% CI [−.10, .18], p = .532) and nor were the effects of warmth at 6 years on child academic test performance at 7 years (β = −.01, 95% CI [−.12, .01], p = .881), child language at 4.5 years on warmth at 6 years (β = −.01, 95% CI [−.16, .15], p = .944), and child language at 6 years on warmth at 7 years (β = −.09, 95% CI [−.22, .03], p = .149). There was no evidence of any indirect effects of birth mother intellectual performance on adoptive parent 2 warmth at 6 or 7 years old via earlier language. Nor was there any evidence of indirect effects of birth mother intellectual performance on child language at 6 years or academic performance at 7 years old via earlier parental warmth.

Figure 4.2 Longitudinal Structural Equation Model Examining the Effects of Birth Mother Intellectual Performance on, and Cross-lagged Associations Between, Adoptive Parent 2 Warmth and Children's Language and Academic Performance



Note. Model fit: $\chi^2(630) = 1284$, $p < .001$, CFI = .88 RMSEA = .05, SRMR = .07. Standardized estimates reported. Solid paths represent significant associations ($p < .05$). Faded paths represent non-significant associations ($p \geq .05$). Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother. AP1 = adoptive parent 1. $^{ns}p \geq .1$. $^{\dagger}\, p < .1$. $^{*}p < .05$. $^{***}p < .001$.

4.3.2 Birth Father Effects on, and Cross-lagged Associations Between, Adoptive Parent Warmth and Children's Language and Academic Performance

Adoptive Parent 1.

Birth Father Effects on Warmth. The significant direct effect in the birth mother model (Figure 4.1) on adoptive parent 1 warmth at 6 years old was not fully reproduced in the semi-independent replication using birth father intellectual ability as a proxy for genetic influence (Figure 4.3). While the effect was in the same (positive) direction and the confidence intervals of the two effects overlapped, the estimate in the birth father model was smaller and not significant at 6 years (β = .07, 95% CI [–.15, .28], *p* = .530). The same is true of the indirect effect on adoptive parent 1 warmth at 7 years via warmth at 6 years (β = .06, 95% CI [–.12, .24], *p* = .530). As in the birth mother model, the direct effect of birth father intellectual performance on adoptive parent 1 warmth was not significant at 4.5 years (β = .10, 95% CI [–.18, .38], *p* = .473), or 7 years (β = .00, 95% CI [–.25, .25], *p* = .988). Although these estimates were not comparable in terms of the direction or effect size to those in the birth mother model, their confidence intervals overlapped, indicating a partial replication of the results. The model accounted for 3% of the variance in adoptive parent 1 warmth at 4.5 years, 72% of the variance in warmth at 6 years and 71% of the variance in warmth at 7 years.

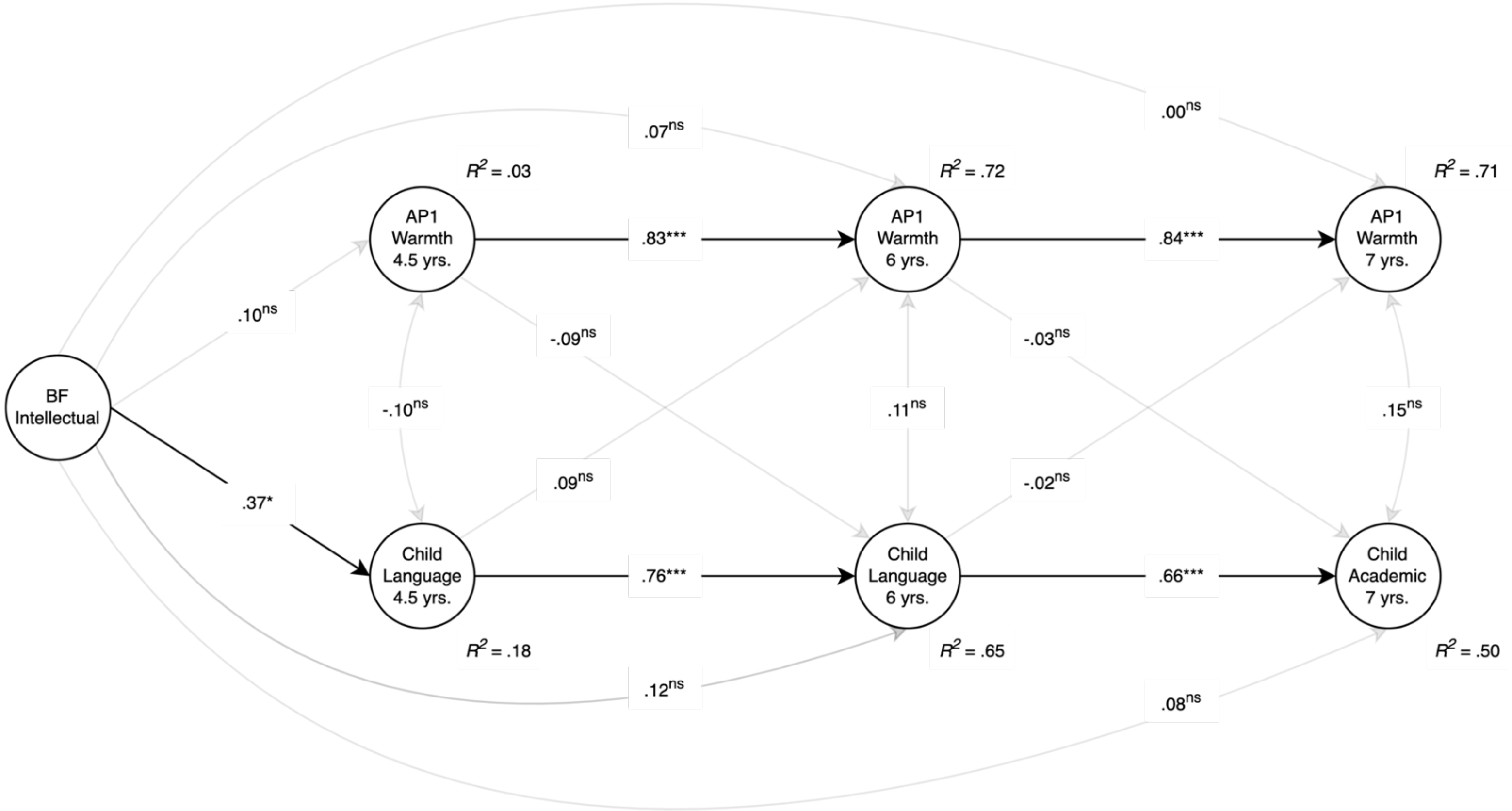
Cross-lagged Associations. The cross-lagged effects in the birth father replication (Figure 4.3) were similar in size and direction to those in the birth mother model (Figure 4.1). In the birth father replication, the direct effect of adoptive parent 1 warmth at 4.5 years on child language at 6 years old was not statistically significant (β = –.09, 95% CI [–.22, .04], *p* = .182) and nor were the effects of warmth at 6 years on child academic test performance at 7 years (β = –.03, 95% CI [–.15, .09], *p* = .636), child language at 4.5 years on warmth at 6 years (β = .09, 95% CI [–.06, .24], *p* = .255), and child language at 6 years on warmth at 7 years (β = –.02, 95% CI [–.17, .13], *p* = .797). As in the birth mother model, there was no evidence of any indirect effects of birth father intellectual performance on adoptive parent 1 warmth at 6 or 7 years old via earlier language. Nor was there any evidence of indirect effects of birth father intellectual performance on child language at 6 years or academic performance at 7 years old via earlier parental warmth.

Adoptive Parent 2.

Birth Father Effects on Warmth. Contrary to the results in the birth mother model (Figure 4.2), in the birth father replication (Figure 4.4), there was a moderate effect in the unexpected (negative) direction for the association between birth father intellectual performance and adoptive parent 2 warmth at 6 years old, although this effect was not significant and its confidence interval overlapped with the confidence interval in the birth mother model suggesting partial replication of results ($\beta = -.22$, 95% CI [−.53, .08], $p = .151$). As in the birth mother model, the direct effect of birth father intellectual performance on adoptive parent 2 warmth was not significant at 4.5 years ($\beta = -.01$, 95% CI [−.30, .28], $p = .949$). When children were 7 years old, the coefficient was of comparable size (but slightly smaller) and in the same direction as the significant association in the birth mother model, with a confidence interval that overlapping with the confidence interval in the birth mother model, however, it was not statistically significant ($\beta = .14$, 95% CI [−.08, .36], $p = .224$). The model accounted for 2% of the variance in adoptive parent 2 warmth at 4.5 years, 69% of the variance in warmth at 6 years and 66% of the variance in warmth at 7 years.

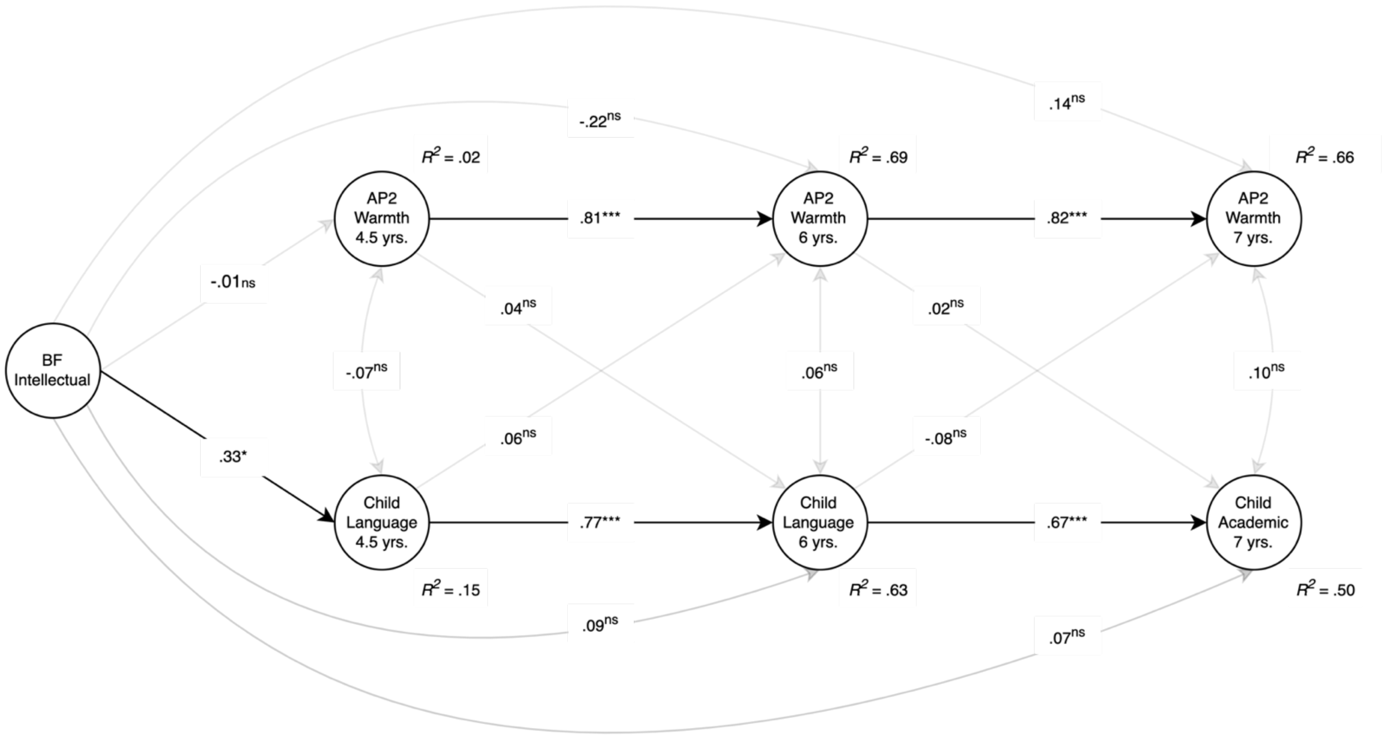
Cross-lagged Associations. As in the birth mother model (Figure 4.2), in the birth father model in Figure 4.4, the direct effect of adoptive parent 2 warmth at 4.5 years on child language at 6 years old was not statistically significant ($\beta = .04$, 95% CI [−.10, .17], $p = .587$) and nor were the effects of warmth at 6 years on child academic test performance at 7 years ($\beta = .02$, 95% CI [−.10, .41], $p = .752$), child language at 4.5 years on warmth at 6 years ($\beta = .06$, 95% CI [−.12, .23], $p = .515$) and child language at 6 years on warmth at 7 years ($\beta = -.08$, 95% CI [−.23, .06], $p = .269$). As in the birth mother model, there was no evidence of any indirect effects of birth father intellectual performance on adoptive parent 2 warmth at 6 or 7 years old via earlier language. Nor was there any evidence of indirect effects of birth father intellectual performance on child language at 6 years or academic performance at 7 years old via earlier parental warm.

Figure 4.3 Longitudinal Structural Equation Model Examining the Effects of Birth Father Intellectual Performance on, and Cross-lagged Associations Between, Adoptive Parent 1 Warmth and Children's Language and Academic Performance



Note. Model fit: $\chi^2(630) = 1373$, $p < .001$, CFI = .86 RMSEA = .05, SRMR = .09. Standardized estimates reported. Solid paths represent significant associations ($p < .05$). Faded paths represent non-significant associations ($p \geq .05$). Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother. AP1 = adoptive parent 1. $^{ns}p \geq .1$. $^{*}p < .05$. $^{***}p < .001$.

Figure 4.4 Longitudinal Structural Equation Model Examining the Effects of Birth Father Intellectual Performance on, and Cross-lagged Associations Between, Adoptive Parent 2 Warmth and Children’s Language and Academic Performance



Note. Model fit: $\chi^2(630) = 1360$, $p < .001$, CFI = .86 RMSEA = .05, SRMR = .08. Standardized estimates reported. Solid paths represent significant associations ($p < .05$). Faded paths represent non-significant associations ($p \geq .05$). Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother. AP1 = adoptive parent 1. $^{ns}p \geq .1$. † $p < .1$. $^{*}p < .05$. $^{***}p < .001$.

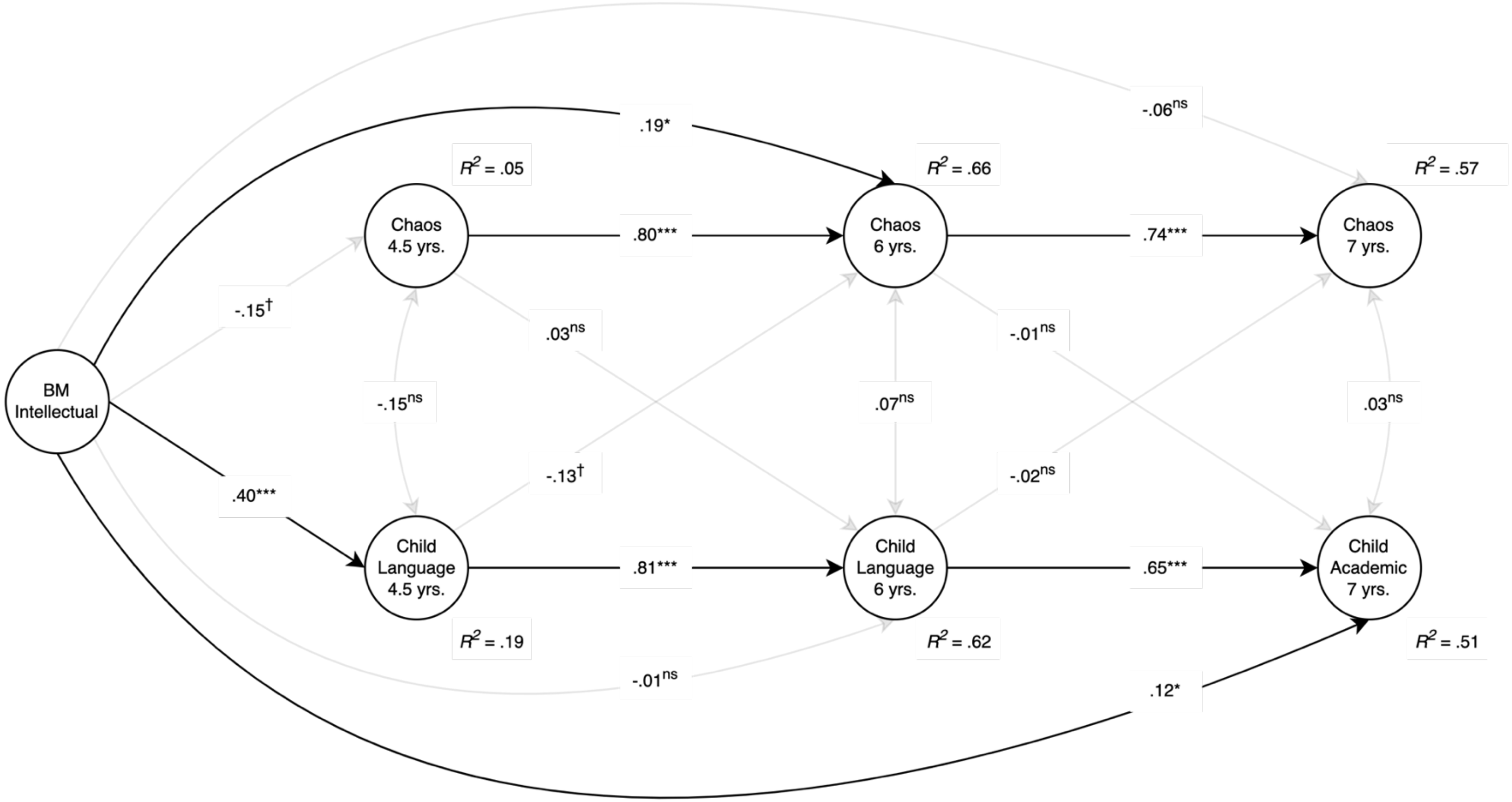
4.3.3 Exploratory Analyses on Birth Parent Effects on, and Cross-lagged Associations Between, Household Chaos and Children's Language and Academic Performance

Birth Mother Effects on Household Chaos. As displayed in Figure 4.5, there was a significant direct effect of birth mother intellectual performance (a proxy for genetic influences) in the unexpected (positive) direction on household chaos at 6 years old (β = .19, 95% CI [.04, .34], p = .012). The direct effect of birth mother intellectual performance on household chaos was in the expected (negative) direction but not statistically significant at 4.5 years (β = –.15, 95% CI [–.31, .01], p = .074) and 7 years (β = –.06, 95% CI [–.21, .09], p = .465). However, the indirect effect of birth mother intellectual performance on chaos at 7 years via chaos at 6 years was statistically significant, again in the unexpected (positive) direction (β = .14, 95% CI [.03, .25], p = .014). The model accounted for 5% of the variance in household chaos at 4.5 years, 66% of the variance in chaos at 6 years and 57% of the variance in chaos at 7 years.

Cross-lagged Associations in Birth Mother Model. As shown in Figure 4.5, the direct effect of household chaos at 4.5 years on child language at 6 years old was not statistically significant (β = .03, 95% CI [–.11, .18], p = .649) and nor were the effects of chaos at 6 years on child academic test performance at 7 years (β = –.01, 95% CI [–.12, .10], p = .852), child language at 4.5 years on chaos at 6 years (β = –.13, 95% CI [–.28, .02], p = .093), and child language at 6 years on chaos at 7 years (β = –.02, 95% CI [–.17, .12], p = .775). There was no evidence of any indirect effects of birth mother intellectual performance on chaos at 6 or 7 years old via earlier language. Nor was there any evidence of indirect effects of birth mother intellectual performance on child language at 6 years or academic performance at 7 years old via earlier chaos.

Birth Father Effects on Household Chaos. When birth father intellectual performance was used as a proxy for genetic influences, instead of birth mother intellectual performance, the model did not converge.

Figure 4.5 Longitudinal Structural Equation Model Examining the Effects of Birth Mother Intellectual Performance on, and Cross-lagged Associations Between, Home Chaos and Children's Language and Academic Performance



Note. Model fit: $\chi^2(342) = 735$, $p < .001$, CFI = .90, RMSEA = .05, SRMR = .07. Standardized estimates reported. Solid paths represent significant associations ($p < .05$). Faded paths represent non-significant associations ($p \geq .05$). Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother. $^{ns}p \geq .1$. † $p < .1$. $^{*}p < .05$. $^{***}p < .001$.

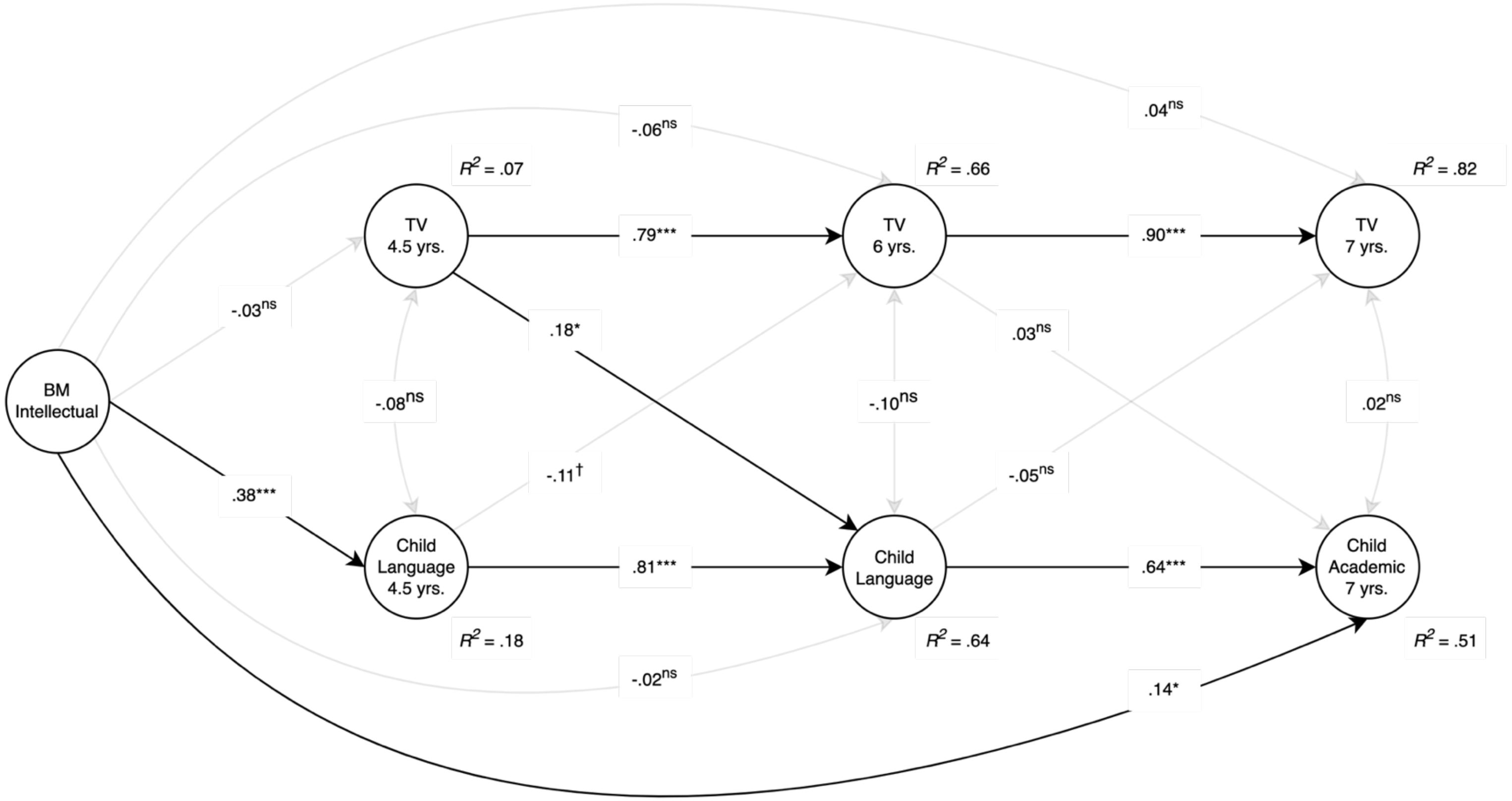
4.3.4 Exploratory Analyses on Birth Parent Effects on, and Cross-lagged Associations Between, Screen Use and Children's Language and Academic Performance

Birth Mother Effects on Screen Use. As displayed in Figure 4.6, the direct effects of birth mother intellectual performance (a proxy for genetic influences) on screen use were small and non-significant at 4.5 years (β = –.03, 95% CI [–.17, .11], p = .697), 6 years old (β = –.06, 95% CI [–.19, .07], p = .381) and 7 years old (β = .04, 95% CI [–.08, .16], p = .507). The model accounted for 7% of the variance in screen use at 4.5 years, 66% of the variance in screen use at 6 years and 82% of the variance in screen use at 7 years.

Cross-lagged Associations in Birth Mother Model. As shown in Figure 4.6, the direct effect of screen use at 4.5 years on child language at 6 years old was in the unexpected (positive) direction and statistically significant (β = .18, 95% CI [.06, .31], p = .005). The effect of screen use at 6 years on child academic test performance at 7 years was not significant (β = .03, 95% CI [–.09, .14], p = .633), nor were the effects of child language at 4.5 years on screen use at 6 years (β = –.12, 95% CI [–.25, .02], p = .083) and child language at 6 years on screen use at 7 years (β = –.05, 95% CI [–.16, .05], p = .321). There was no evidence of any indirect effects of birth mother intellectual performance on screen use at 6 or 7 years old via earlier language. Nor was there any evidence of indirect effects of birth mother intellectual performance on child language at 6 years or academic performance at 7 years old via earlier household screen use.

Birth Father Effects on Screen Use. When birth father intellectual performance was used as a proxy for genetic influences, instead of birth mother intellectual performance, the model would not converge.

Figure 4.6 Longitudinal Structural Equation Model Examining the Effects of Birth Mother Intellectual Performance on, and Cross-lagged Associations Between, Screen use and Children's Language and Academic Performance



Note. Model fit: $\chi^2(429) = 1209$, $p < .001$, CFI = .83, RMSEA = .07, SRMR = .08. Standardized estimates reported. Solid paths represent significant associations ($p < .05$). Faded paths represent non-significant associations ($p \geq .05$). Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother. TV = screen use (television watching and video gaming). $^{ns}p \geq .1$. † $p < .1$. $^{*}p < .05$. $^{***}p < .001$.

4.4 Discussion

This study leveraged the parent-offspring adoption design to test for evocative gene-environment correlation (*r*GE) in the development of children's academic abilities. The results demonstrated that birth mother intellectual performance (used as a proxy for genetic influence on children's academic outcomes) predicted adoptive parent 1 warmth when children were 6 years and 7 years old, but not 4.5 years old, and adoptive parent 2 warmth at 7 years, but not at 4.5 or 6 years. These findings are partially consistent with the first study hypothesis (that there would be evocative effects on parenting of genetic influences underlying children's academic ability) and evidence from twin and polygenic score research on evocative *r*GE in early cognitive development (Tucker-Drob & Harden, 2012; Wertz et al., 2020). However, results from the birth mother models were not fully replicated in the analysis using data from birth fathers as the proxy for genetic influences. Although the effects that were significant in the birth mother models were generally in the same (positive) direction as in the birth father models and had overlapping confidence intervals, the effect sizes were smaller and not statistically significant in the birth father models.

Contrary to the second and third hypotheses that evocative effects on parenting would mediate genetic effects on academic outcomes and that language would be a mechanism through which genetic influences would evoke parenting differences), there was no evidence of bidirectional associations between adoptive parent warmth and children's language and academic outcomes or indirect genetic effects via parenting or child language. Although these findings indicate that adoptive parents might parent their children differently depending on their children's academically associated genes, they do not demonstrate a mediating influence of these parenting differences on children's academic test performance in middle childhood. The only statistically significant effects on children's language and academic test outcomes were the direct effects of birth parent intellectual performance reported in the earlier work in **Chapter 3** (Austerberry, Fearon, et al., 2022). Furthermore, despite this earlier work indicating that language appears to be an early manifestation of genetic influences on later academic outcomes, the present findings do not indicate that early language evokes differences in parental warmth.

Consequently, the present findings failed to uncover evidence to support the Dickens and Flynn (2001) hypothesis that increasing heritability of intellectual performance across the lifespan is produced through a process of amplification via mechanisms of reciprocal causation between an individual's genotype and the environmental influences it evokes over time.

It is surprising, based on the well-established associations between positive parenting and children's intellectual and educational outcomes (Fan & Chen, 2001; Lugo-Gil & Tamis-LeMonda, 2008; Madigan et al., 2019; Wertz et al., 2020), that parental warmth did not predict children's language or academic performance. This lack of association would suggest that previous findings may be attributable to passive rGE. Indeed, Wertz et al. (2020) demonstrated in the Environmental Risk (E-Risk) Study, firstly, that parent and child education polygenic scores were each pleiotropic: associated with parental warmth *and* children's educational performance. Secondly, when they partialled out the effects of child polygenic scores, the association reduced in size, indicating some genetic confounding. Although, in contrast to the present findings, the association between parental warmth and educational achievement still held after the polygenic score was partialled out. However, as Wertz et al. (2020) discuss (because of the limitations of polygenic scores discussed in **Chapter 1** and below), unlike the adoption design, their methods do not rule out the possibility of passive rGE and they likely only control for a proportion of the genetic confounding. Another possible explanation for the lack of association between parental warmth and educational performance is that warmth may be less important for educational performance than other dimensions of positive parenting. Indeed, Lugo-Gil and Tamis-LeMonda (2008) and Wertz et al. (2020) both used more global measures of positive parenting, the former combining measures of maternal supportiveness, sensitivity, positive regard and cognitive stimulation and the latter combining measures of warmth, sensitivity and reverse coded negative parenting. Furthermore, the meta-analysis of parenting and academic achievement by Fan and Chen (2001) specifically found that parental involvement was associated with children's academic achievement ($\bar{r} = .30$), and in the series of meta-analyses by Madigan et al. (2019), the pooled effect was larger for the association between sensitive-responsive parenting and child language ($r = 0.27$) than for the association between parental warmth and language ($r = 0.16$). The null findings in the present analyses may also reflect issues with measurement or limitations to generalisability of the present findings due to use of an unrepresentative sample (discussed, in the limitations section, below).

In exploratory analyses of home chaos and screen media use, the only aspect of the home environment significantly predicted by differences in genetic influences underlying children's educational outcomes was home chaos. Contrary to the broader literature, which has tended to report negative associations between household chaos and educational performance (Johnson et al., 2008; Petrill et al., 2004), the effect was positive: higher birth mother intellectual performance directly predicted higher home chaos when children were 6 years old and indirectly predicted chaos at 7 years old via chaos at 6 years old. There was only one significant cross-lagged association in the exploratory analyses: more screen media use (a composite of items on television watching and video gaming) at 4.5 years old was associated with higher language performance at 6 years old. While this finding is in line with results from some studies (Adelantado-Renau et al., 2019), it runs counter to the more widely reported finding that screen media use in childhood and adolescence is negatively associated with academic performance (Adelantado-Renau et al., 2019). This may reflect the possibility that a high SES sample of parents who have chosen to adopt children are more likely to monitor the types of screen time their children are exposed to (e.g., the programs their children watch) than a representative sample would. Finally, although the present findings indicate that screen media use in the home may positively influence language, they do not suggest that screen media use in the home mediates genetic influences on language.

4.4.1 Limitations and Future Directions

This research is the first (to my knowledge) to examine evocative *r*GE in intellectual development using data from an adoption study, providing a powerful control against passive *r*GE. The assumption that passive *r*GE is ruled out depends on the environments of adoptees not being influenced by their birth parents. Early placement of EGDS adoptees (on average 6 days postpartum) reduces the likelihood of this assumption being violated. However, there are two potential threats to this presumption that are important to consider. First, any ongoing contact between adoptees and their birth parents introduces the possibility of birth parents influencing adoptees' environments. I attempted to control this potential confound by including a composite of birth and adoptive parent ratings of adoption openness as a covariate in the analyses. The second potential threat is that passive *r*GE could occur if any aspects of the prenatal environment are correlated with the genes that birth parents pass on to their children. I controlled for this by including a measure of prenatal risk as a covariate in the models and replicating the analyses in the birth father sample. Birth father analyses are a robust control for prenatal effects only if birth fathers do not indirectly affect the prenatal environment by, for example, contributing to home dynamics or stress levels of the mother. The low rates of birth mother cohabitation (6%) in the EGDS sample plausibly reduce the chances of indirect effects of birth fathers on the foetal environment.

Overall, it is a strength of the present study that I was able to run two sets of analyses: the main analysis using a sample of birth mothers and a quasi-independent replication of this analysis using a sample of birth fathers. This is a particular strength given that fathers are so under-researched relative to mothers in developmental research. However, the birth father analyses are not fully independent as most measures of birth mother and birth father intellectual performance were associated, suggesting the possibility of assortative mating, confounding, or partner interaction effects (Austerberry, Fearon, et al., 2022). Despite being the largest sample of birth fathers ever recruited in a prospective parent-offspring adoption study, the birth father analyses were limited by their small sample size, resulting in a lack of statistical power to accurately estimate the influence of birth father intellectual performance. This lack of power may explain why the findings from the birth mother models were not fully replicated in the birth father analyses. However, it is impossible to rule out the possibility that lack of robust replication signals either spurious results in the birth mother models or prenatal (rather than genetic) effects being detected in associations between birth mother intellectual performance and adoptive parent parenting. Although the likelihood of the latter is probably reduced by the inclusion of a prenatal risk covariate, replication of these methods using larger samples of birth fathers is needed before decisive conclusions can be drawn.

Although the adoption design has notable strengths, an inherent limitation is that there is relatively little variation in family income and parental education in adoptive families relative to the general population. The adoptive families in the EGDS sample have a higher socioeconomic status (SES) than the birth parents in the study and the general population of the US, potentially biasing the results (Leve, Neiderhiser, et al., 2013a). Thus, it remains to be seen whether the present findings would replicate in lower SES, more ethnically/racially diverse families, particularly as SES appears to moderate genetic effects on educational outcomes (Capron & Duyme, 1989; Tucker-Drob & Bates, 2015) and may causally influence parenting (Akee et al., 2010). Research suggests that adoptees in the US and UK perform better than expected academically (based on their preadoption intelligence scores, education polygenic scores or comparisons with their biological relatives) after being adopted (Cheesman et al., 2020; Duyme et al., 1999; Kendler et al., 2015), indicating mediation or moderation of genetic effects on academic performance by differences in caregiving environments (or the wider social conditions that are associated with them). There is also evidence from the US and UK that trajectories of language and academic development are not the same for different ethnic groups, and this may be explained by differences in psychosocial, family and home environments (Saccuzzo et al., 1992; Zilanawala et al., 2016). Although almost half of the adoptees in the EGDS sample were multiracial, Black/African American, or Latinx, over 90% of the adoptive parents in the study were non-Latinx White and the study was US-based, adding to the literature on samples from Western, educated, industrialized, rich and democratic populations, who, despite making up approximately 12% of the world's population, are the subject of the vast majority of findings published in top psychology research journals (Arnett, 2008). Replication of the present methods in different populations is needed to address this stark inequity and before it can be assumed that the results generalize.

Finally, there are three potential limitations concerning measurement and the operationalization of analysed constructs. First, there may be limits to the extent to which the measures analysed were suitable for testing the Dickens and Flynn (2001) hypothesis. I focused my analyses primarily on parental warmth, on the basis that prior research had demonstrated not only that parental positivity and warmth appear to predict better cognitive and educational outcomes in children but also that children's cognitive abilities, or genetic propensities linked to cognitive and educational abilities, appear to positively predict the warmth and positivity of the parenting they receive (Lugo-Gil & Tamis-LeMonda, 2008; Madigan et al., 2019; Wertz et al., 2020). However, another dimension of parenting that prior evidence suggests may be evoked by children's cognition and education linked genetic propensities is cognitively stimulating parenting (Tucker-Drob & Harden, 2012; Wertz et al., 2020), which (as I discussed in detail above) may also be important for children's educational development. There were not suitable data available in the pre-existing dataset used for the present analyses to test my hypothesis using measures of parental cognitive stimulation. However, it would be of interest for future work to incorporate such measures. Furthermore, Dickens and Flynn (2001) describe a multiplicity of many individual and social factors over time (including activities undertaken during leisure time, cognitive quality of social interactions, cognitive demands at school, and cognitive complexity of work) not necessarily each having a large effect but cumulatively resulting in substantial change in cognitive ability across development. Consequently, a fairer test of the Dickens and Flynn (2001) hypothesis would incorporate many varied measures over time, rather than focus specifically on parenting during a narrow age range. Second, as the present results rely on self-reports by parents of their parenting and the caregiving environment, they are vulnerable to reporter bias and ceiling effects, which occur when a large proportion of respondents score near the upper limit of a scale so that variance is not measured above a certain level. In the present sample, most adoptive parents rated their parenting as 'almost always' or 'always' warm, resulting in low variability in the responses. Consequently, future replication should incorporate observational measurement. Third, I did not use direct genetic measures and instead relied on measures of birth parent traits as indirect proxies for the genetic load. However (as discussed at length in **Chapter 1**), it remains open for debate which behavior genetic methods best capture the full contribution of genetic influences, as there is a discrepancy (known as "missing" heritability) between estimates from genome-wide analyses and those relying on family data such as adoption or, more commonly, twin studies. For example, in the most recently published genome-wide association study of educational attainment, a polygenic index explained 12-16% of the variance in educational attainment (Okbay et al., 2022)—around one-third of the size of the 43% heritability estimate reported in a recent analysis of a pooled sample of 28 twin cohorts (Silventoinen et al., 2020). Consequently, it remains helpful to continue triangulating findings from studies using different methods. The present analyses contributed to this effort as evocative *r*GE in intellectual development had (to my knowledge) only ever been examined previously using the twin design and polygenic score analyses, making this study the first to examine this mechanism in academic development using an adoption study and birth parent trait status as a proxy for genetic influences.

4.4.2 Conclusion

In this first (to my knowledge) empirical test of the Dickens and Flynn (2001) hypothesis and the first examination of evocative *r*GE in academic development in an adoption sample, I found some evidence of evocative effects of genetic influences underlying children's academic development on parental warmth. While these effects did not seem to be evoked by differences in early language and did not mediate associations between genetic influences and middle-childhood academic outcomes, they nonetheless converge with findings from twin and polygenic score research in suggesting that parents may parent their children differently depending on children's genetic predispositions for academic attainment.

4.5 References

Adelantado-Renau, M., Moliner-Urdiales, D., Cavero-Redondo, I., Beltran-Valls, M. R., Martínez-Vizcaíno, V., & Álvarez-Bueno, C. (2019). Association between screen media use and academic performance among children and adolescents: A systematic review and meta-analysis. *JAMA Pediatrics, 173*(11), 1058-1067. https://doi.org/10.1001/jamapediatrics.2019.3176

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

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *The Journal of Abnormal Psychology, 112*, 545-557. https://doi.org/10.1037/0021-843X.112.4.545

Arnett, J. J. (2008, Oct). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602-614. https://doi.org/10.1037/0003-066x.63.7.602

Austerberry, C., & Fearon, P. (2021). An overview of developmental behavioral genetics. In L. Provenzi & R. Montirosso (Eds.), *Developmental human behavioral epigenetics* (Vol. 23, pp. 59-80). Academic Press. https://doi.org/10.1016/B978-0-12-819262-7.00003-9

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., & Reiss, D. (2021, March 24). Evocative effects of genetic influences underlying the development of intellectual ability on the early caregiving environment. *OSF Preregistration*. https://doi.org/10.17605/OSF.IO/PZ67M

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., & Reiss, D. (2022). Early manifestations of intellectual performance: Evidence that genetic effects on later academic test performance are mediated through verbal performance in early childhood. *Child Development, 93*(2), e188-e206. https://doi.org/10.1111/cdev.13706

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75*(2), 81-95. https://doi.org/10.1037/h0025583

Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*(1), 78-117. https://doi.org/10.1177/0049124187016001004

Blair, C., Raver, C. C., Berry, D. J., & Family Life Project Investigators. (2014, Feb). Two approaches to estimating the effect of parenting on the development of executive function in early childhood. *Developmental Psychology, 50*(2), 554-565. https://doi.org/10.1037/a0033647

Boivin, M., Pérusse, D., Dionne, G., Saysset, V., Zoccolillo, M., Tarabulsy, G. M., Tremblay, N., & Tremblay, R. E. (2005). The genetic-environmental etiology of parents' perceptions and self-assessed behaviours toward their 5-month-old infants in a large twin and singleton sample. *Journal of Child Psychology and Psychiatry, 46*(6), 612-630. https://doi.org/https://doi.org/10.1111/j.1469-7610.2004.00375.x

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20*, 115-140. https://doi.org/10.2307/271084

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*(4498), 1055-1059. https://doi.org/10.1126/science.7195071

Capron, C., & Duyme, M. (1989). How much can we raise or lower I.Q.? The joint effects genetic and prenatal factors versus postnatal factors. *Journal de Genetic Human, 37*(2), 115-118.

Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., Eley, T. C., & Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological Science, 31*(5), 582-591. https://doi.org/10.1177/0956797620904450

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346. https://doi.org/10.1037/0033-295X.108.2.346

Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of “dull children”?: A late adoption study. *Proceedings of the National Academy of Sciences, 96*(15), 8790-8794. https://doi.org/10.1073/pnas.96.15.8790

Elam, K. K., Harold, G. T., Neiderhiser, J. M., Reiss, D., Shaw, D. S., Natsuaki, M. N., Gaysina, D., Barrett, D., & Leve, L. D. (2014). Adoptive parent hostility and children's peer behavior problems: Examining the role of genetically informed child attributes on adoptive parent behavior. *Developmental Psychology, 50*(5), 1543-1552. https://doi.org/10.1037/a0035470

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1-22. https://doi.org/10.1023/A:1009048817385

Fearon, P., Reiss, D., Leve, L. D., Shaw, D. S., Scaramella, L. V., Ganiban, J. M., & Neiderhiser, J. M. (2015). Child-evoked maternal negativity from 9 to 27 months: Evidence of gene–environment correlation and its moderation by marital distress. *Development and Psychopathology, 27*(4pt1), 1251-1265. https://doi.org/10.1017/S0954579414000868

Ge, X., Natsuaki, M. N., Martin, D. M., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Villareal, G., Scaramella, L., Reid, J. B., & Reiss, D. (2008). Bridging the divide: Openness in adoption and postadoption psychosocial adjustment among birth and adoptive parents. *The Journal of Family Psychology, 22*, 529-540. https://doi.org/10.1037/a0012817

Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Institute for the Development of Educational Achievement.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015, Mar). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102-116. https://doi.org/10.1037/a0038889

Harold, G. T., Leve, L. D., Barrett, D., Elam, K. K., Neiderhiser, J. M., Natsuaki, M. N., Shaw, D. S., Reiss, D., & Thapar, A. (2013). Biological and rearing mother influences on child ADHD symptoms: Revisiting the developmental interface between nature and nurture. *Journal of Child Psychology and Psychiatry, 54*(10), 1038-1046. https://doi.org/10.1111/jcpp.12100

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S. P., Kovas, Y., Corley, R. P., DeFries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*(11), 1112-1120. https://doi.org/10.1038/mp.2009.55

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

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.

Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55. https://doi.org/10.1080/10705519909540118

Hummer, R. A., & Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States. *Population bulletin, 68*(1), 1-16.

Johnson, A. D., Martin, A., Brooks-Gunn, J., & Petrill, S. A. (2008). Order in the House! Associations among Household Chaos, the Home Literacy Environment, Maternal Reading Ability, and Children's Early Reading. *Merrill Palmer Q (Wayne State Univ Press), 54*(4), 445-472. https://doi.org/10.1353/mpq.0.0009

Kendler, K. S., Turkheimer, E., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proceedings of the National Academy of Sciences, 112*(15), 4612-4617. https://doi.org/10.1073/pnas.1417106112

Klahr, A. M., Burt, S. A., Leve, L. D., Shaw, D. S., Ganiban, J. M., Reiss, D., & Neiderhiser, J. M. (2017). Birth and adoptive parent antisocial behavior and parenting: A study of evocative gene–environment correlation. *Child Development, 88*(2), 505-513. https://doi.org/10.1111/cdev.12619

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. The Guilford Press.

Knafo, A., & Plomin, R. (2006). Parental discipline and affection and children's prosocial behavior: Genetic and environmental links. *Journal of Personality and Social Psychology, 90*(1), 147. https://doi.org/10.1037/0022-3514.90.1.147

Kosik, R., Mandell, G., Fan, A. P., Nguyen, T., Chen, J., & Eaton, W. (2018). The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study. *The European Journal of Psychiatry, 32*(2), 53-62. https://doi.org/10.1016/j.ejpsy.2018.01.001

Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*(10), 2048-2056. https://doi.org/10.1177/0956797613486982

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 22*(6), 716-727. https://doi.org/10.1017/thg.2019.66

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J., Natsuaki, M. N., & Reiss, D. (2013, Feb). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*(1), 412–423. https://doi.org/https://doi.org/10.1017/thg.2012.126

Lonigan, C. J., Allan, N. P., & Lerner, M. D. (2011). Assessment of preschool early literacy skills: Linking children's educational needs with empirically supported instructional activities. *Psychology in the Schools, 48*, 488-501. https://doi.org/10.1002/pits.20569

Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children’s cognitive development across the first 3 years. *Child Development, 79*(4), 1065-1085. https://doi.org/10.1111/j.1467-8624.2008.01176.x

Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., & Jenkins, J. M. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics, 144*(4). https://doi.org/10.1542/peds.2018-3556

Marceau, K., De Araujo-Greecher, M., Miller, E. S., Massey, S. H., Mayes, L. C., Ganiban, J. M., Reiss, D., Shaw, D. S., Leve, L. D., & Neiderhiser, J. M. (2016). The perinatal risk index: Early risks experienced by domestic adoptees in the United States. *PLoS One, 11*, e0150486. https://doi.org/10.1371/journal.pone.0150486

Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of Applied Developmental Psychology, 16*(3), 429-444. https://doi.org/10.1016/0193-3973(95)90028-4

McNeil, T. F., Cantor-Graae, E., & Sjostrom, K. (1994). Obstetric complications as antecedents of schizophrenia: Empirical effects of using different obstetric complication scales. *Journal of Psychiatric Research, 28*, 519-530. https://doi.org/10.1016/0022-3956(94)90042-6

Melby, J. N., & Conger, R. D. (2001). The Iowa Family Interaction Rating Scales: Instrument summary. In P. K. Kerig & K. M. Lindahl (Eds.), *Family observational coding systems: Resources for systemic research* (pp. 33–58). Erlbaum.

Niklas, F., & Schneider, W. (2013). Home literacy environment and the beginning of reading and spelling. *Contemporary Educational Psychology, 31*(1), 40-50.

Nunnally, J. C. (1967). *Psychometric theory*. McGraw-Hill.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlskog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., Porteous, D. J., Freese, J., Herd, P., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., Hinds, D. A., Huber, K. E., Kleinman, A., Litterman, N. K., McCreight, J. C., McIntyre, M. H., Mountain, J. L., Northover, C. A. M., Pitts, S. J., Sathirapongsasuti, J. F., Sazonova, O. V., Shelton, J. F., Shringarpure, S., Tung, J. Y., Vacic, V., Wilson, C. H., Fontana, M. A., Pers, T. H., Rietveld, C. A., Chen, G.-B., Emilsson, V., Meddens, S. F. W., Pickrell, J. K., Thom, K., Timshel, P., de Vlaming, R., Abdellaoui, A., Ahluwalia, T. S., Bacelis, J., Baumbach, C., Bjornsdottir, G., Brandsma, J. H., Concas, M. P., Derringer, J., Galesloot, T. E., Girotto, G., Gupta, R., Hall, L. M., Harris, S. E., Hofer, E., Horikoshi, M., Huffman, J. E., Kaasik, K., Kalafati, I. P., Karlsson, R., Lahti, J., van der Lee, S. J., de Leeuw, C., Lind, P. A., Lindgren, K.-O., Liu, T., Mangino, M., Marten, J., Mihailov, E., Miller, M. B., van der Most, P. J., Oldmeadow, C., Payton, A., Pervjakova, N., Peyrot, W. J., Qian, Y., Raitakari, O., Rueedi, R., Salvi, E., Schmidt, B., Schraut, K. E., Shi, J., Smith, A. V., Poot, R. A., Pourcain, B. S., Teumer, A., Thorleifsson, G., Verweij, N., Vuckovic, D., Wellmann, J., Westra, H.-J., Yang, J., Zhao, W., Zhu, Z., Alizadeh, B. Z., Amin, N., Bakshi, A., Baumeister, S. E., Biino, G., Bønnelykke, K., Boyle, P. A., Campbell, H., Cappuccio, F. P., Davies, G., De Neve, J.-E., Deloukas, P., Demuth, I., Ding, J., Eibich, P., Eisele, L., Eklund, N., Evans, D. M., Faul, J. D., Feitosa, M. F., Forstner, A. J., Gandin, I., Gunnarsson, B., Halldórsson, B. V., Harris, T. B., Heath, A. C., Hocking, L. J., Holliday, E. G., Homuth, G., Horan, M. A., Hottenga, J.-J., de Jager, P. L., Joshi, P. K., Jugessur, A., Kaakinen, M. A., Kähönen, M., Kanoni, S., Keltigangas-Järvinen, L., Kiemeney, L. A. L. M., Kolcic, I., Koskinen, S., Kraja, A. T., Kroh, M., Kutalik, Z., Latvala, A., Launer, L. J., Lebreton, M. P., Levinson, D. F., Lichtenstein, P., Lichtner, P., Liewald, D. C. M., Loukola, A., Madden, P. A., Mägi, R., Mäki-Opas, T., Marioni, R. E., Marques-Vidal, P., Meddens, G. A., McMahon, G., Meisinger, C., Meitinger, T., Milaneschi, Y., Milani, L., Montgomery, G. W., Myhre, R., Nelson, C. P., Nyholt, D. R., Ollier, W. E. R., Palotie, A., Paternoster, L., Pedersen, N. L., Petrovic, K. E., Räikkönen, K., Ring, S. M., Robino, A., Rostapshova, O., Rudan, I., Rustichini, A., Salomaa, V., Sanders, A. R., Sarin, A.-P., Schmidt, H., Scott, R. J., Smith, B. H., Smith, J. A., Staessen, J. A., Steinhagen-Thiessen, E., Strauch, K., Terracciano, A., Tobin, M. D., Ulivi, S., Vaccargiu, S., Quaye, L., van Rooij, F. J. A., Venturini, C., Vinkhuyzen, A. A. E., Völker, U., Völzke, H., Vonk, J. M., Vozzi, D., Waage, J., Ware, E. B., Willemsen, G., Attia, J. R., Bennett, D. A., Berger, K., Bertram, L., Bisgaard, H., Boomsma, D. I., Borecki, I. B., Bültmann, U., Chabris, C. F., Cucca, F., Cusi, D., Deary, I. J., Dedoussis, G. V., van Duijn, C. M., Eriksson, J. G., Franke, B., Franke, L., Gasparini, P., Gejman, P. V., Gieger, C., Grabe, H.-J., Gratten, J., Groenen, P. J. F., Gudnason, V., van der Harst, P., Hoffmann, W., Hyppönen, E., Iacono, W. G., Jacobsson, B., Järvelin, M.-R., Jöckel, K.-H., Kaprio, J., Kardia, S. L. R., Lehtimäki, T., Lehrer, S. F., Martin, N. G., McGue, M., Metspalu, A., Pendleton, N., Penninx, B. W. J. H., Perola, M., Pirastu, N., Pirastu, M., Polasek, O., Posthuma, D., Power, C., Province, M. A., Samani, N. J., Schlessinger, D., Schmidt, R., Sørensen, T. I. A., Spector, T. D., Stefansson, K., Thorsteinsdottir, U., Thurik, A. R., Timpson, N. J., Tiemeier, H., Uitterlinden, A. G., Vitart, V., Vollenweider, P., Weir, D. R., Wilson, J. F., Wright, A. F., Conley, D. C., Krueger, R. F., Smith, G. D., Hofman, A., Laibson, D. I., Medland, S. E., Yang, J., Esko, T., Watson, C., Jala, J., Conley, D., Koellinger, P. D., Johannesson, M., Laibson, D., Meyer, M. N., Lee, J. J., Kong, A., Yengo, L., Cesarini, D., Turley, P., Visscher, P. M., Beauchamp, J. P., andMe Research, T., Social Science Genetic Association, C., & LifeLines Cohort, S. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics, 54*(4), 437-449. https://doi.org/10.1038/s41588-022-01016-z

Pardini, D. A., Fite, P. J., & Burke, J. D. (2008). Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: The moderating effect of age and african-american ethnicity. *Journal of Abnormal Child Psychology, 36*(5), 647-662. https://doi.org/10.1007/s10802-007-9162-z

Petrill, S. A., Pike, A., Price, T., & Plomin, R. (2004). Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. *Intelligence, 32*, 445–460.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on “environmental” measures. *Behavioral and Brain Sciences, 14*(3), 373-386. https://doi.org/10.1017/S0140525X00070278

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.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/https://doi.org/10.18637/jss.v048.io2

Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal versus performance IQs for gifted African-American, Caucasian, Filipino, and Hispanic children. *Psychological Assessment, 4*(2), 239. https://doi.org/10.1037/1040-3590.4.2.239

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424-435. https://doi.org/10.2307/1129703

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., Tomizawa, R., Watanabe, M., Sakai, N., Rebato, E., Busjahn, A., Tyler, J., Hopper, J. L., Ordoñana, J. R., Sánchez-Romera, J. F., Colodro-Conde, L., Calais-Ferreira, L., Oliveira, V. C., Ferreira, P. H., Medda, E., Nisticò, L., Toccaceli, V., Derom, C. A., Vlietinck, R. F., Loos, R. J. F., Siribaddana, S. H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Duncan, G. E., Buchwald, D., Tynelius, P., Rasmussen, F., Tan, Q., Zhang, D., Pang, Z., Magnusson, P. K. E., Pedersen, N. L., Dahl Aslan, A. K., Hwang, A. E., Mack, T. M., Krueger, R. F., McGue, M., Pahlen, S., Brandt, I., Nilsen, T. S., Harris, J. R., Martin, N. G., Medland, S. E., Montgomery, G. W., Willemsen, G., Bartels, M., van Beijsterveldt, C. E. M., Franz, C. E., Kremen, W. S., Lyons, M. J., Silberg, J. L., Maes, H. H., Kandler, C., Nelson, T. L., Whitfield, K. E., Corley, R. P., Huibregtse, B. M., Gatz, M., Butler, D. A., Tarnoki, A. D., Tarnoki, D. L., Park, H. A., Lee, J., Lee, S. J., Sung, J., Yokoyama, Y., Sørensen, T. I. A., Boomsma, D. I., & Kaprio, J. (2020). Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. *Scientific Reports, 10*(1), 12681. https://doi.org/10.1038/s41598-020-69526-6

Taylor, L. C., Clayton, J. D., & Rowley, S. J. (2004). Academic socialization: Understanding parental influences on children's school-related development in the early years. *Review of General Psychology, 8*(3), 163-178. https://doi.org/10.1037/1089-2680.8.3.163

Tucker-Drob, E. M., & Bates, T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138-149. https://doi.org/10.1177/0956797615612727

Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene–environment transactions. *Developmental Science, 15*(2), 250-259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. Psychological Corporation.

Wechsler, D. (2002). *WPPSI-III administration and scoring manual.* Psychological Corporation.

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd,

L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745-1761. https://doi.org/10.1111/cdev.13329

Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson tests of achievement (WJ III®)*. Riverside Publishing.

Zilanawala, A., Kelly, Y., & Sacker, A. (2016). Ethnic differences in longitudinal latent verbal profiles in the millennium cohort study. *European Journal of Public Health, 26*(6), 1011-1016. https://doi.org/10.1093/eurpub/ckw184

Chapter 5: Examination of Evocative Effects on Parenting of Common Genetic Variants Associated with Educational Attainment

5.1 Introduction

As discussed in **Chapters 3 and 4** (Austerberry, Fearon, et al., 2022), Intellectual and academic abilities are important assets in their own right and powerful predictors of health and longevity (Deary et al., 2010; Hummer & Hernandez, 2013; Kosik et al., 2018). Understanding their aetiology is important and has the potential to uncover modifiable mechanisms implicated in their development. Intellectual and academic performance are highly heritable, especially with age, rising from approximately 20–50% in childhood and adolescence to around 50–80% in adulthood (Bouchard & McGue, 1981; C. M. A. Haworth et al., 2010; Kovas et al., 2013). Although this might be interpreted as suggesting that environmental factors play only a minor role in their longitudinal development, a plausible alternative hypothesis (as discussed in **Chapter 4** of this thesis) is that these increasing heritability estimates mask key environmental influences. Specifically, it has been hypothesised that the environment may have a mediating, or even amplifying, effect on genetic influences through mechanisms of gene-environment correlation (*r*GE) (Dickens & Flynn, 2001; Plomin et al., 1977; Scarr & McCartney, 1983). Evocative *r*GE, which may be the most important form of *r*GE in early childhood (when children have limited opportunities for active selection of their environment), occurs when an individual's genetically influenced characteristics systematically evoke differences in their environment (e.g., the early caregiving environment), mediating genetic effects. As evoked environmental differences would be correlated with genetic differences, their effects could be masked by global estimates of genetic effects.

There have been many recent studies, precipitated by the landmark study by Kong et al. (2018) examining 'genetic nurture', investigating the indirect influence of parents' genes on children's educational outcomes via environmental mechanisms (Wang et al., 2021). However, there has been far less research on the indirect influence of children's genes on their educational outcomes via environmental mechanisms such as evocative *r*GE. There is good reason to believe that children's early characteristics can elicit responses in their caregivers, evidenced by a robust body of literature (reviewed in **Chapter**

4) that supports the existence of interplay between child traits and the parenting environment. However, as summarised below (and reiterating **Chapter 4**), only a small subset of this literature is specifically focused on intellectual and academic development.

There is evidence of evocative *r*GE in early verbal and non-verbal cognitive development from one study that used multivariate twin modelling (Tucker-Drob & Harden, 2012). Specifically, Tucker-Drob and Harden (2012) found that, after controlling for earlier cognitively stimulating parenting, children's cognitive performance at 2 years of age predicted cognitively stimulating parenting at 4 years old and the association was almost entirely genetically mediated. Chapter 4 of this thesis also found that genetic influences underlying children's educational performance (measured in an adoption sample using birth parent academic test performance as a proxy) positively predicted positive parenting and negatively predicted household chaos (when children were between 5 to 10 years of age), indicating possible evocative *r*GE, although these effects did not predict subsequent child academic test performance. Genomic studies on evocative *r*GE are scarce and I am aware of only two previous studies that have used genomic data to examine associations between children's education-associated genes and the early caregiving environment (Krapohl et al., 2017; Wertz et al., 2020). The first of these studies found children's education polygenic scores were associated with aspects of their early caregiving, such as length of time they were breastfed for, parental smacking they experienced, and number of books in the home (Krapohl et al., 2017). However, as parental genotypes were not controlled for, it was not possible to rule out passive *r*GE in these associations. The second study found that children's education polygenic scores were positively associated with positive and cognitively stimulating maternal parenting and household chaos, after controlling for mothers', but not fathers', polygenic scores (Wertz et al., 2020). To my knowledge, evocative effects on parenting of education polygenic scores have never been examined while controlling for polygenic scores from both parents. I aimed to do this using data from the Norwegian Mother, Father and Child Cohort Study (MoBa). I also aimed to examine whether any observed evocative effects on parenting would predict children's subsequent academic performance. Wider research, including **Chapter 3** of this thesis, suggests that early childhood language may be an early manifestation of genetic influences on later cognitive and academic performance (Austerberry, Fearon, et al., 2022; Verhoef et al., 2021) and may evoke differences in parenting that influence children's academic outcomes (Tucker-Drob & Harden, 2012). Guided by this evidence, I also investigated whether early childhood language mediated any observed genetic effects on later academic performance and was associated with early parenting.

Specifically, I examined the following hypotheses. First, that common genetic variants associated with educational attainment would have evocative effects on maternal parenting in early childhood. Specifically, I expected that children's educational attainment polygenic scores would predict maternal parenting when children were 5 years old, after controlling for mothers' and fathers' education polygenic scores. Second, that any observed evocative effects on maternal parenting would predict subsequent child academic performance. Specifically, I expected for parenting, when children were 5 years old, to mediate the expected positive association between children's education polygenic scores and their academic performance in grades 1 and 2 (when they were 6–8 years old). Third, that language skills at 5 years of age would mediate genetic effects on academic performance in grades 1 and 2 and be associated with parenting when children were 5 years old. These hypotheses are important to examine because they may uncover mechanisms that are potentially implicated in the causal chain from genotype to intellectual and academic outcomes, aiding research into promotion and prevention efforts.

5.2 Methods

5.2.1 Sample

Data were from the Norwegian Mother, Father and Child Cohort Study (MoBa), a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health (Magnus et al., 2016; Magnus et al., 2006). Pregnant women were recruited from all over Norway from 1999 to 2008. The women consented to participation in 41% of the pregnancies (N = 112,908 recruited pregnancies). The cohort now includes 114,500 children, 95,200 mothers and 75,200 fathers. The current study is based on version 12 of the quality-assured data files released for research in January 2019. The establishment of MoBa and initial data collection was based on a license from the Norwegian Data protection agency and approval from The Regional Committees for Medical and Health Research Ethics. The MoBa cohort is based on regulations based on the Norwegian Health Registry Act. The current study was approved by The Regional Committees for Medical and Health Research Ethics (21076). Blood samples were collected from both parents during pregnancy and from children (umbilical cord) at birth (Paltiel et al., 2014). Further information on recruitment and data collection has been reported in published cohort profiles for the study (Magnus et al., 2016; Magnus et al., 2006). Protocols (which include consent forms and questionnaires) can be found here: https://www.fhi.no/en/publ/2012/protocols-for-moba/. The MoBa assessments used in the current analyses include questionnaires sent to mothers when children were 5 and 8 years old. I also used genetic data from a subsample of 98,110 genotyped individuals in mother-father-child trios.

The socioeconomic status of individuals included in the main analyses was relatively high compared to the general population of women and men in Norway during the study recruitment period (1999–2008). For example, 67% of mothers included in the main analyses, and 52% of included fathers, had completed tertiary education, compared to 21% of women and 21% of men in 1999 (Statistisk Sentralbyrå, 2001), and 29% of women and 25% of men in 2008 (Statistisk Sentralbyrå, 2009).

5.2.2 Genotyping and Genotype Quality Control

The genotyping has been performed in the following three projects (and some smaller projects): HARVEST (~33,000 individuals), ERC HARVEST (~27,000 individuals) NORMENT (> 100,000 individuals). The present analysis used data (version 1.0) released by MoBa genetics in 2019 (https://github.com/folkehelseinstituttet/mobagen). This data includes 98,110 samples genotyped in ten batches. After release, the data went through genotype quality control (QC; by colleagues and I, between March 2020 and September 2020) imputation and post-imputation QC (by colleagues) using the family based MoBa PsychGen pipeline (Corfield et al., 2022). The pre-imputation quality control (QC) and imputation were performed in line with current best-practice QC protocol (https://github.com/Nealelab/picopili/blob/master/bin/imp_prep.pl). The primary software used for the QC was PLINK 1.9 (Purcell et al., 2007) and KING 2.2.5

(Manichaikul et al., 2010). To identify core continental subpopulations (European, African, and Asian), principal component analyses (PCA) with 1000 Genomes reference were performed. Prior to the PCA, SNPs with minor allele frequency (MAF) below 1%, call rate below 95% and out of Hardy-Weinberg Equilibrium (HWE, $p < 0.001$), as well as individuals with call rate below 95%, were removed. Pre-imputation QC was performed for each of the core subpopulations on the SNP and individual level. QC on a SNP level involved filtering for 0.5% MAF, 95% call rate, HWE *p*-value 0.000001, discordance in duplicate pairs, association with genotype plate and genotype batch at *p*-value of 0.001. Individual level QC was performed by filtering for heterozygosity outliers (Fhet ± 0.2), erroneous sex assignment, known relatedness, cryptic relatedness (PI_HAT > 15% in unrelated individuals), and outliers in PCA both with and without 1000 Genomes. Mendelian errors (which occur when alleles are found in an offspring that could not have been obtained through Mendelian inheritance from either of their parents) were assessed for families with a minimum of one parent-offspring (PO) duo. Families with more than 5% Mendelian errors and SNPs with more than 1% of Mendelian errors were removed, while other minor Mendelian errors were set to zero. Batches that were genotyped using versions of the same array were merged (keeping only SNPs present in all batches per merge) and the pre-imputation QC was performed on the three merged batches (OMNI, GSA, and HCE). Phasing and imputation were performed using the publicly available Haplotype Reference Consortium data. Phasing was performed using SHAPEIT2 with the duoHMM algorithm to incorporate the pedigree information into the haplotype estimates. IMPUTE 4 was then used to perform imputation. For post-imputation QC, dosage data of SNPs with imputation quality score (INFO) of 80% and above were converted to best-guess genotypes, using default PLINK certainty of 90%. Post-imputation QC was then performed following the steps outlined in the pre-imputation QC. The integrity of relatedness across the batches (both known, such as PO and full sibling [FS] relationships, and unknown, such as sibships within the parent generation) was ensured in all analyses. The three imputation batches were merged, and post-imputation QC was performed on the overall merged dataset.

Ahead of the present analysis, colleagues and I performed additional post-imputation QC on each imputation batch, separately. Imputed best-guessed genotype data were subjected to post-imputation QC performed in PLINK 1.9 (Purcell et al., 2007), during which we removed SNPs that had call rate below 98%, had MAF less than 1%, were out of HWE ($p < 1.00 \times 10^{-06}$) and showed significant difference in their MAFs between the genotyping batches ($p < 1.00 \times 10^{-03}$). Further, we removed individuals who had call rate below 98% and/or showed too much or too little heterozygosity (inbreeding coefficient more or less than 0.20). Relatedness was examined in KING 2.2.5 (Manichaikul et al., 2010) and further confirmed in PLINK 1.9 (Purcell et al., 2007) with PI_HAT estimates. Unrelated individuals were ensured to have PI_HAT measure below 20% by removing one individual from each pair of unrelated individuals who showed more than 20% PI_HAT, prioritizing the inclusion of children, then mothers and then fathers. An additional relatedness check was performed across all batches combined to ensure that there was no cryptic relatedness. Mendelian errors were checked in PLINK 1.9 (Purcell et al., 2007) and families with more than 5% errors as well as SNPs with more than 1% of errors were removed. All remaining Mendelian errors were set to missing. European ancestry of participants was ensured by performing PCA in PLINK 1.9 (Purcell et al., 2007) of MoBa samples together with reference populations from the 1,000 genomes project (McVean et al., 2012) and removing individuals clustering outside European reference populations. To examine any possible technical artifacts, we performed PCA in PLINK 1.9 (Purcell et al., 2007) of MoBa samples only. The first ten principal components from these analyses were consequently included as covariates in the analyses. These QC steps were performed in each imputation batch separately. Next, trios and duos that were imputed in different batches were combined from each imputation batch and QCed as one batch (CROSS batch), resulting in four batches after the QC had been completed: (1) OMNI batch: 19,302 individuals and 4,982,332 SNPs, (2) GSA batch: 30,507 individuals and 4,988,017 SNPs, (3) HCE batch: 24,686 individuals and 4,974,358 SNPs, and (4) CROSS batch: 1,967 individuals and 4,769,645 SNPs.

5.2.3 Measures

Education Polygenic Scores. Genetic predisposition for educational attainment was measured using years of education (EduYears) polygenic scores. I constructed the polygenic scores using PRS-CS, which is a Python based command line tool that infers posterior SNP effect sizes under continuous shrinkage priors (Ge et al., 2019). I calculated EduYears polygenic scores for individuals (children, mothers and fathers) using publicly available summary statistics (not including 23andme) from the Lee et al. (2018) EduYears GWAS (known as EA3) and a European LD reference panel constructed using the 1000 genomes project phase 3 samples (McVean et al., 2012). First, I matched the EA3 summary statistics (10,101,242 SNPs) to the LD reference panel (1,120,696 SNPs), resulting in 1,110,307 overlapping SNPs. Next, I used PRS-CS to shrink the SNP effect sizes (standardized regression coefficients), using the EA3 summary statistics, the reference panel and an artificially constructed bim file that I created, containing all overlapping SNPs. I used the following PRS-CS parameters: 25,000 Markov chain Monte Carlo (MCMC) iterations, 10,000 burnin iterations, global shrinkage factor = 0.02, parameter a in the gamma-gamma prior = 1, parameter b in the gamma-gamma prior = 0.5, MCMC thinning factor = 5. Second, I QCed the 1,110,307 overlapping SNPs in the summary statistics in line with Choi et al. (2020). Specifically, I checked for and, if identified, removed: variants that were not SNPs or were strand ambiguous, indels, SNPs with missing values, SNPs in the summary statistics with MAF less than 1%, SNPs in the LD reference panel with MAF less than 1%, SNPs with MAF difference between summary statistics and reference panel greater than 20%, mismatching alleles between the summary statistics and reference panel, SNPs with out of bounds p-values, and SNPs with duplicated rs numbers. The files outputted by PRS-CS containing the shrunk regression coefficients were then matched with the 1,105,512 remaining EA3 SNPs. Third, I prepared the target data in line with Choi et al. (2020). Specifically, in each of the four batches of (post-QC) MoBa genetic data I selected only common SNPs (with MAF less than 1%) and those with INFO 90% or above. Each batch was then matched to the LD reference data and SNPs were removed that had MAF difference between the target and reference data of greater than 20%. I then checked if the minor alleles matched between the target and reference data and removed mismatching SNPs. The number of variants and individuals remaining in each batch were: (1) OMNI batch: 19,302 individuals and 739,509 SNPs, (2) GSA batch: 30,507 individuals and 739,699 SNPs, (3) HCE batch: 24,686 individuals and 738,413 SNPs and (4) CROSS batch: 1,967 individuals and 726,108 SNPs. Finally, the polygenic scores were calculated in PLINK 1.9 (Purcell et al., 2007), for each batch separately, using the PRS-CS output (after it had been matched with the post-QC summary statistics SNPs). Polygenic scores were calculated for 76,462 individuals (25,623 children, 26,024 mothers, 24815 fathers) from 36,468 trios (OMNI batch: 10,563 trios, GSA batch: 13,804 trios, HCE batch: 11,349 trios, CROSS batch: 752 trios). Prior to hypothesis testing, the polygenic scores were standardised to have a mean of 0 and variance of 1. Residualised polygenic scores were then calculated by regressing the standardised polygenic scores on genotyping plate (to control for plate effects, which have been detected in the MoBa sample) and the first 10 principal components from a PCA (to control for population stratification). These standardised and residualised polygenic scores were used in the analyses. I included all trios in the main analyses that had genetic data for at least one member of the trio (36,468 trios).

Maternal Parenting.

Alabama Parenting Questionnaire (APQ). Using confirmatory factor analysis (CFA), I constructed a latent variable using six items from the ‘Positive Parenting’ subscale of the APQ, a 42-item scale developed by Frick (1991) to assess parenting practices. The six items were included in the questionnaire sent out to mothers when the children were 5 years old, and all answers were scored on a 5-point scale from ‘1-Never’ to ‘5-Always’. Research has established good validity and reliability of the APQ (Essau et al., 2006; Frick et al., 1999; Shelton et al., 1996b). In the sample retained for the main analyses, reliability was good ($\alpha = .77$).

Cognitively Stimulating Parenting. Two items from the questionnaire sent to mothers when the children were 5 years old were averaged to create a total score. The two questions were selected by MoBa from the Early Language in Victoria Study (ELVS) (Prior et al., 2011): (1) ‘During a typical week, how often do you teach your child how to print letters and words?’; (2) ‘During a typical week, how often do you help your child read letters and sounds?’. Both items were scored on a 5-point Likert from ‘1-Never’ to ‘5-Very Often’. In the sample included in the main analyses, the two items were highly correlated (r = .69, $p < .001$).

Child Language.

Speech and Language Assessment Scale (SLAS). Using CFA, I constructed a latent variable with indicators from the SLAS, which was included in the questionnaire sent to mothers when children were 5 years old. The SLAS is a reliable and well validated 14-item measure aiming to capture children's articulation, semantics, vocabulary, sentence construction and conversational skills compared with peers (Hadley & Rice, 1993; Rice et al., 1989). All items were scored on a 5-point Likert from '1-Very much lower' to '5-Very much higher'. In the sample used for the main analyses, the fourteen items had high reliability (α = .96).

Child Academic Performance. I constructed a latent variable (using CFA) from three items in the questionnaire sent out to mothers when the children were 8 years old about the feedback they had been given during parent-teacher discussions concerning their child's performance in national exams on: (1) '… Reading skills in 1st grade', (2) 'Reading skills in 2nd grade', (3) 'Arithmetic skills in 2nd grade'. Each item was scored on a 3-point Likert: '1-Has mastered subject well', '2-Must work more but teacher is not concerned', '3-Teacher is concerned'. In the sample used for the main analysis, reliability was good (α = .75). Prior to constructing the latent variable, the items were reverse scored so that higher scores represented higher reading and arithmetic skills.

Covariates. As well as residualising the polygenic scores by controlling for genotyping plate and the top ten principal components, child sex and year of birth were included as covariates in the main analyses.

5.2.4 Statistical Analysis

The analyses were preregistered with the Open Science Framework in February 2022 (Austerberry, Zayats, et al., 2022). The hypotheses were tested in the lavaan package version 0.6-10 (Rosseel, 2012) in R version 4.0.3 (R Core Team, 2020) using structural equation modelling (SEM), which combines latent variable models (CFAs) with structural models examining hypothesized causal relations (also known as a path models). I constructed the SEMs in several steps: First, I ran bivariate SEMs, examining associations between

EduYears polygenic scores (child, maternal, and paternal, separately) and the four phenotypic variables: (1) latent variable measuring child academic performance at 6–8 years, (2) latent variable measuring child language at 5 years, (3) latent variable measuring maternal positive parenting at 5 years, (4) maternal cognitively stimulating parenting at 5 years old. Second, I ran four trio SEMs, which jointly modelled the effects of child, mother, and father polygenic scores on the four aforementioned phenotypic variables. The two trio SEMs examining the effects of trio polygenic scores on maternal parenting at 5 years old were a test of study hypothesis 1: that there would be evocative effects of children's education polygenic scores on parenting, after controlling for parent polygenic scores. Third, I examined whether the effect of children's education-associated genes on their academic performance at 6–8 years old would be mediated via parenting at 5 years old (study hypothesis 2) in two models (one examining mediation via positive maternal parenting and the second testing for mediation via cognitively stimulating maternal parenting). Specifically, I combined each of the two trio SEMs examining the effects of trio polygenic scores on parenting with the trio SEM examining the effects of trio polygenic scores on academic performance at 6–8 years old and calculated the mediated effect of children's polygenic scores on their academic performance via maternal parenting. Fourth, I examined whether the effect of children's education-associated genes on their academic performance at 6–8 years old was mediated via their language at 5 years old (study hypothesis 2). Specifically, I combined the trio SEM examining the effects of the trio polygenic scores on language with each of the two models (one on positive maternal parenting and the second on cognitively stimulating maternal parenting) constructed in the aforementioned third step (immediately above) and calculated the mediated effect of children's polygenic scores on their academic performance at 6–8 years old via their language performance at 5 years old. I also included a path in the model, examining the association between maternal parenting at 5 years old and children's language at 5 years old. The mediated effects were estimated using bootstrapping with 5000 repetitions (Bollen & Stine, 1990).

I used $p < .05$ as the criteria for determining whether associations in the SEMs were statistically significant. Based on recommendations by Hu and

Bentler (1999), I used a combination rule, according to which model fit was considered adequate if SRMR < .09 and RMSEA < .06. All trios that had genotypic data available for at least one member of the mother-father-child trio were included in the analyses and families with no genotypic data were excluded from the analyses. I tested for multicollinearity by running separate multiple regression models based on estimated factor scores saved from each of the structural equation models that jointly modelled child, mother, and father EduYears polygenic scores. VIF > 4 were interpreted as indicating the presence of multicollinearity. In the main analyses, missing data in the structural equation models was handled using full information maximum likelihood (FIML), which simulation studies suggest outperforms listwise deletion and produces unbiased parameter estimates and standard errors when the data are missing at random (MAR) or missing completely at random (MCAR) (Enders & Bandalos, 2001). The data used in the study were not missing completely at random (MCAR) according to results from the Little's MCAR test, conducted in the naniar package version 0.6.1 (Tierney et al., 2021) in R version 4.0.3 (R Core Team, 2020): χ^2 (6178 = 7292, $p < .001$. Consequently, I ran an additional attrition analysis using the Missing Value Analysis function in IBM SPSS Statistics for Windows version 28.0, which creates an indicator variable identifying variables that contain missing values. This indicator value was then used to compare group means among different variables in the dataset, using the *t*-test procedure. In the attrition analysis, the patterns of missingness for all study variables were related to the observed values of one or more other variables in the dataset. These results ruled out the possibility that the data were MCAR, which occurs when the probability of being missing is the same for all cases and there is no systematic association between the missingness of the data and any other values, observed or missing. It was not possible to rule out the possibility that the data were missing not at random (MNAR), which is when the missingness of the data is systematically related to unobserved data. However, the results from the attrition analysis were consistent with the data being missing at random (MAR), which occurs when the missingness of a variable is systematically related to the observed but not the unobserved data.

5.3 Results

5.3.1 Descriptive Statistics

Means, standard deviations, sample sizes and bivariate correlations between study variables are presented in Table 5.1.

5.3.2 Effects of Child, Mother, and Father Polygenic Scores on Study Variables

Prior to hypothesis testing, I ran bivariate structural equation models (SEMs), examining the associations between each of the three EduYears polygenic scores (child, mother, and father), individually, and the four variables included in the main models: (1) children's academic performance at 6–8 years old; (2) children's language at 5 years old, (3) positive maternal parenting at 5 years old, (4) cognitively stimulating parenting at 5 years old. The results from these bivariate SEMs are presented in Table 5.2. Next, I ran trio SEMs, jointly modelling the effects of the three EduYears polygenic scores on each of the four variables. The results from this second set of analyses are reported in Figure 5.1 and outlined, below.

Effects of Polygenic Scores on Child Academic Performance and Language.

Effects of Polygenic Scores on Academic Performance. As in the bivariate models reported in Table 5.2, in the trio model presented in Figure 5.1a, there was a statistically significant unmediated effect of children's EduYears polygenic scores on their academic performance at 6–8 years old ($\beta = .10$, 95% CI [.07, .13], $p < .001$). The model in Figure 5.1a, also controlled for parental genetic effects on academic performance at 6–8 years (i.e., the effects originating in the parts of the parental genome not transmitted to the child, commonly known as genetic nurture, represented by the unmediated paths from parental polygenic scores to child academic performance), which were not statistically significant (EduYears$_{\text{mother}}$: $\beta = -.01$, 95% CI [–.03, .02], $p = .718$; EduYears$_{\text{father}}$: $\beta = -.02$, 95% CI [–.05, .01], $p = .290$).

Effects of Polygenic Scores on Language. As in the bivariate models reported in Table 5.2, in the trio model presented in Figure 5.1b, there was a statistically significant unmediated effect of children's EduYears polygenic scores on their language at 5 years old ($\beta = .04$, 95% CI [.01, .07], $p = .013$). As depicted in Figure 5.1b, parental genetic effects on language were not significant (EduYears$_{\text{mother}}$: $\beta = .00$, 95% CI [–.03, .03], $p = .880$; EduYears$_{\text{father}}$: $\beta = .01$, 95% CI [–.02, .04], $p = .528$).

Table 5.1 Means, Standard Deviations and Bivariate Correlations Between Study Variables

Variable	M	SD	n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1. EduYears$_{child}$	0.00	1.00	25623	–																								
2. EduYears$_{mother}$	0.00	1.00	26024	.54	–																							
3. EduYears$_{mother}$	0.00	1.00	24815	.54	.11	–																						
4. APQ (Item 1)	4.45	0.53	7031	-.03	-.06	-.03	–																					
5. APQ (Item 2)	4.19	0.46	7028	-.02	-.05	-.04	.29	–																				
6. APQ (Item 3)	4.78	0.45	7002	-.03	-.05	-.02	.21	.20	–																			
7. APQ (Item 4)	4.64	0.49	7021	-.04	-.06	-.04	.47	.28	.40	–																		
8. APQ (Item 5)	4.45	0.57	7017	-.06	-.09	-.05	.46	.28	.30	.63	–																	
9. APQ (Item 6)	4.23	0.58	7013	-.02	-.05	-.02	.32	.37	.32	.41	.48	–																
10. Cog parenting	3.41	0.84	11481	.00	-.04	.01	.16	.16	.13	.17	.17	.22	–															
11. SLAS (Item 1)	3.44	0.68	11474	.04	.03	.03	.06	.10	.07	.08	.07	.10	.15	–														
12. SLAS (Item 2)	3.52	0.68	11480	.05	.02	.04	.07	.12	.08	.09	.07	.12	.15	.84	–													
13. SLAS (Item 3)	3.62	0.78	11476	.02	.02	.01	.06	.10	.08	.09	.08	.10	.14	.69	.70	–												
14. SLAS (Item 4)	3.73	0.75	11461	.05	.02	.03	.07	.09	.08	.09	.07	.10	.16	.64	.64	.67	–											
15. SLAS (Item 5)	3.61	0.71	11462	.05	.03	.03	.08	.10	.08	.09	.08	.11	.15	.66	.67	.68	.79	–										
16. SLAS (Item 6)	3.57	0.71	11471	.04	.02	.02	.06	.10	.06	.09	.07	.11	.14	.69	.70	.72	.69	.75	–									
17. SLAS (Item 7)	3.52	0.69	11471	.04	.01	.03	.07	.11	.07	.09	.09	.12	.15	.68	.70	.69	.70	.76	.82	–								
18. SLAS (Item 8)	3.54	0.69	11465	.02	.02	.01	.07	.11	.07	.09	.08	.12	.14	.65	.65	.65	.62	.67	.74	.75	–							
19. SLAS (Item 9)	3.47	0.69	11465	.01	.01	.01	.07	.12	.06	.09	.11	.14	.11	.54	.54	.53	.48	.52	.59	.58	.64	–						
20. SLAS (Item 10)	3.44	0.66	11463	.01	.01	.00	.07	.12	.07	.08	.11	.14	.12	.57	.58	.55	.51	.55	.61	.61	.65	.85	–					
21. SLAS (Item 11)	3.63	0.73	11464	.05	.02	.04	.07	.11	.09	.08	.08	.11	.16	.67	.67	.68	.73	.72	.72	.72	.67	.57	.62	–				
22. SLAS (Item 12)	3.59	0.75	11472	.04	.01	.03	.08	.10	.09	.09	.08	.11	.17	.63	.63	.63	.70	.69	.67	.68	.62	.50	.54	.78	–			
23. SLAS (Item 13)	3.44	0.87	11475	.03	.01	.01	.04	.09	.06	.07	.06	.10	.14	.54	.54	.68	.56	.60	.60	.60	.55	.44	.46	.61	.59	–		
24. Reading, Grade 1	2.64	0.54	11428	.08	.04	.04	.01	.01	.02	.01	-.01	.04	.16	.18	.19	.19	.19	.17	.18	.19	.16	.10	.11	.19	.19	.19	–	
25. Reading, Grade 2	2.70	0.52	11280	.07	.03	.03	.02	.02	.02	.02	.00	.04	.14	.16	.17	.17	.17	.17	.17	.18	.15	.09	.10	.18	.17	.17	.76	–
26. Maths, Grade 2	2.82	0.42	11191	.10	.05	.05	.01	.02	.01	-.01	-.01	.01	.06	.13	.13	.12	.14	.14	.14	.14	.12	.07	.08	.15	.15	.11	.36	.39

█	$p < .001$
▓	$p < .01$
▒	$p < .05$
░	$p < .1$
□	$p > .1$

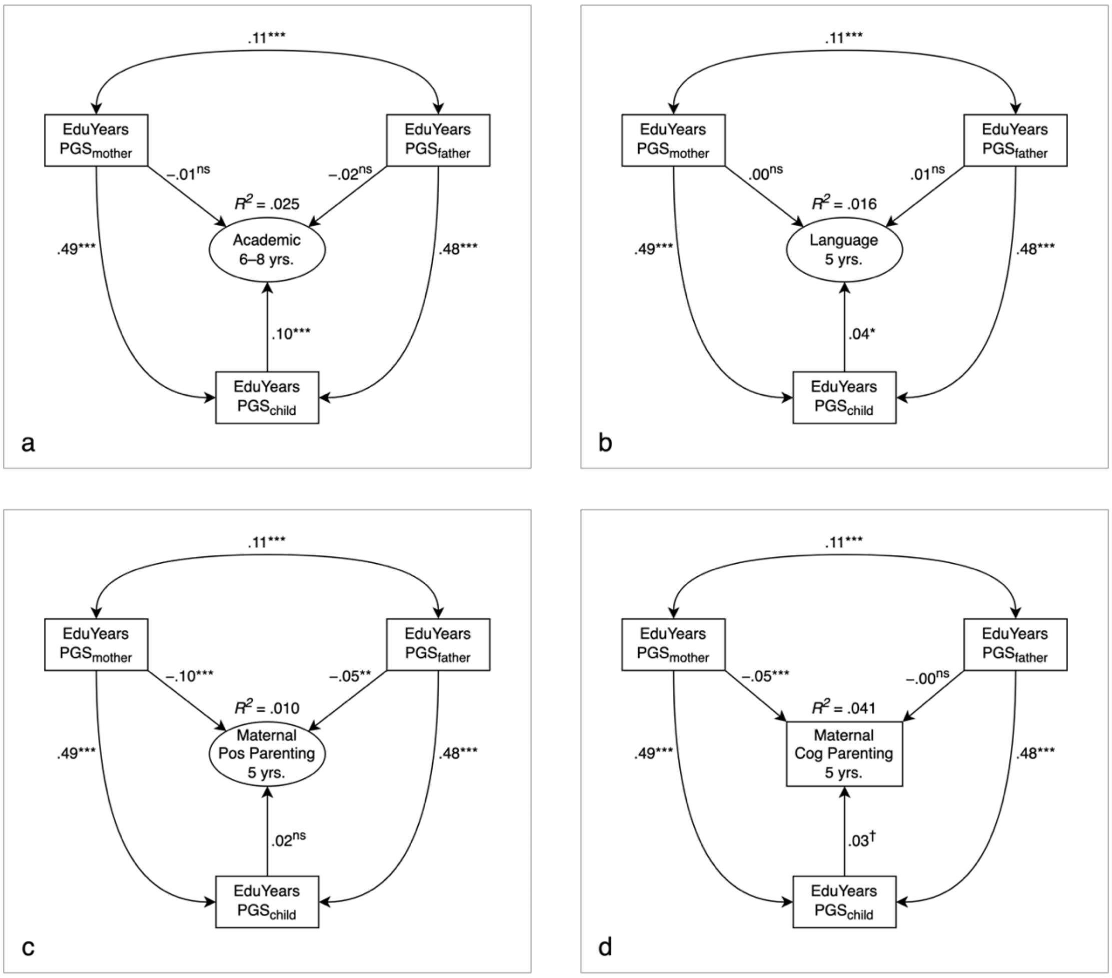
Note. M = Mean. SD = Standard deviation. n = Sample size. EduYears = Years of education polygenic score. APQ = Alabama Parenting Questionnaire. Cog = Cognitively stimulating. SLAS = Speech and Language Assessment Scale.

Table 5.2 Bivariate Structural Equation Models Examining Associations Between Polygenic Scores and Study Variables

Hypothesised Association	β	CI $_{95\%\ lower}$	CI $_{95\%\ upper}$	SE	*p*-value	R^2
Child EduYears PGS → Academic Performance 6–8 yrs.	.09	.07	.11	.01	<.001	.008
Mother EduYears PGS → Academic Performance 6–8 yrs.	.04	.02	.07	.01	<.001	.002
Father EduYears PGS → Academic Performance 6–8 yrs.	.04	.01	.06	.01	.004	.001
Child EduYears PGS → Language 5 yrs.	.05	.03	.06	.01	<.001	.002
Mother EduYears PGS → Language 5 yrs.	.02	-.00	.04	.01	.056	.000
Father EduYears PGS → Language 5 yrs.	.03	.01	.05	.01	.012	.001
Child EduYears PGS → Maternal Positive Parenting 5 yrs.	-.06	-.09	-.03	.01	<.001	.003
Mother EduYears PGS → Maternal Positive Parenting 5 yrs.	-.10	-.13	-.07	.02	<.001	.009
Father EduYears PGS → Maternal Positive Parenting 5 yrs.	-.06	-.09	-.02	.02	<.001	.003
Child EduYears PGS → Maternal Cognitively Stimulating Parenting 5 yrs.	-.00	-.02	.02	.01	.870	.000
Mother EduYears PGS → Maternal Cognitively Stimulating Parenting 5 yrs.	-.04	-.06	-.02	.01	.001	.001
Father EduYears PGS → Maternal Cognitively Stimulating Parenting 5 yrs.	.01	-.02	.03	.01	.531	.000

Note. β = Standardised beta. CI = Confidence interval. SE = Standard error. R^2 = proportion of variance in the dependent variable explained by the independent variable. EduYears PGS = Years of education polygenic score. Standardised estimates reported.

Figure 5.1 Preliminary Trio Models Examining the Effects of Polygenic Scores on Study Variables



Note. EduYears PGS = years of education polygenic score. Pos = Positive. Cog = Cognitively stimulating. Figure 5.1a fit: $\chi^2(16) = 231$, $p < .001$, comparative fit index (CFI) = .99, root mean square error of approximation (RMSEA) = .02, standardized root mean square residual (SRMR) = .02. Figure 5.1b fit: $\chi^2(131) = 19056$, $p < .001$, CFI = .88, RMSEA = .06, SRMR = .04. Figure 5.1c fit: $\chi^2(40) = 743$, $p < .001$, CFI = .97, RMSEA = .03, SRMR = .03. Figure 5.1d fit: $\chi^2(6) = 17$, $p < .001$, CFI = .99, RMSEA = .01, SRMR = .01. Standardised estimates reported. Child sex and year of birth were included as covariates in the model. $^{ns}p \geq .1$; $^{\dagger}p < .1$ $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Test of Hypothesis 1: Evocative Effects of Children’s Polygenic Scores on Parenting, Controlling for Parent Polygenic Scores.

Positive Maternal Parenting. There was no evidence of evocative effects of children’s education-associated genes on maternal positive parenting. While, in the bivariate models reported in Table 5.2, children’s EduYears polygenic scores were associated (in the unexpected, negative, direction) with maternal positive parenting, in the trio model presented in Figure 5.1c, which controlled for parent polygenic scores, the unmediated effect of children’s EduYears polygenic scores on maternal positive parenting at 5 years old was not statistically significant ($\beta = .02$, 95% CI [−.02, .06], $p = .337$). As in the bivariate models in Table 5.2, the unmediated effects of both parental polygenic scores on maternal positive parenting were in the unexpected (negative) direction and statistically significant ($\text{EduYears}_{\text{mother}}$: $\beta = -.10$, 95% CI [−.13, −.06], $p < .001$; $\text{EduYears}_{\text{father}}$: $\beta = -.05$, 95% CI [−.09, −.01], $p = .009$).

Cognitively Stimulating Maternal Parenting. There was no evidence of evocative effects of children’s education-associated genes on maternal cognitively stimulating parenting. As in the bivariate models reported in Table 5.2, in the trio model presented in Figure 5.1d, controlling for parent polygenic scores, the unmediated effect of children’s EduYears polygenic scores on maternal cognitively stimulating parenting at 5 years old was not statistically significant ($\beta = .03$, 95% CI [−.00, .06], $p = .086$). As in the bivariate models in Table 5.2, the unmediated effect of mother polygenic scores, but not father polygenic scores, was statistically significant ($\text{EduYears}_{\text{mother}}$: $\beta = -.05$, 95% CI [−.08, −.02], $p < .001$; $\text{EduYears}_{\text{father}}$: $\beta = -.00$, 95% CI [−.03, .02], $p = .791$). The effect of the mothers’ polygenic scores on their parenting was in the unexpected, negative, direction.

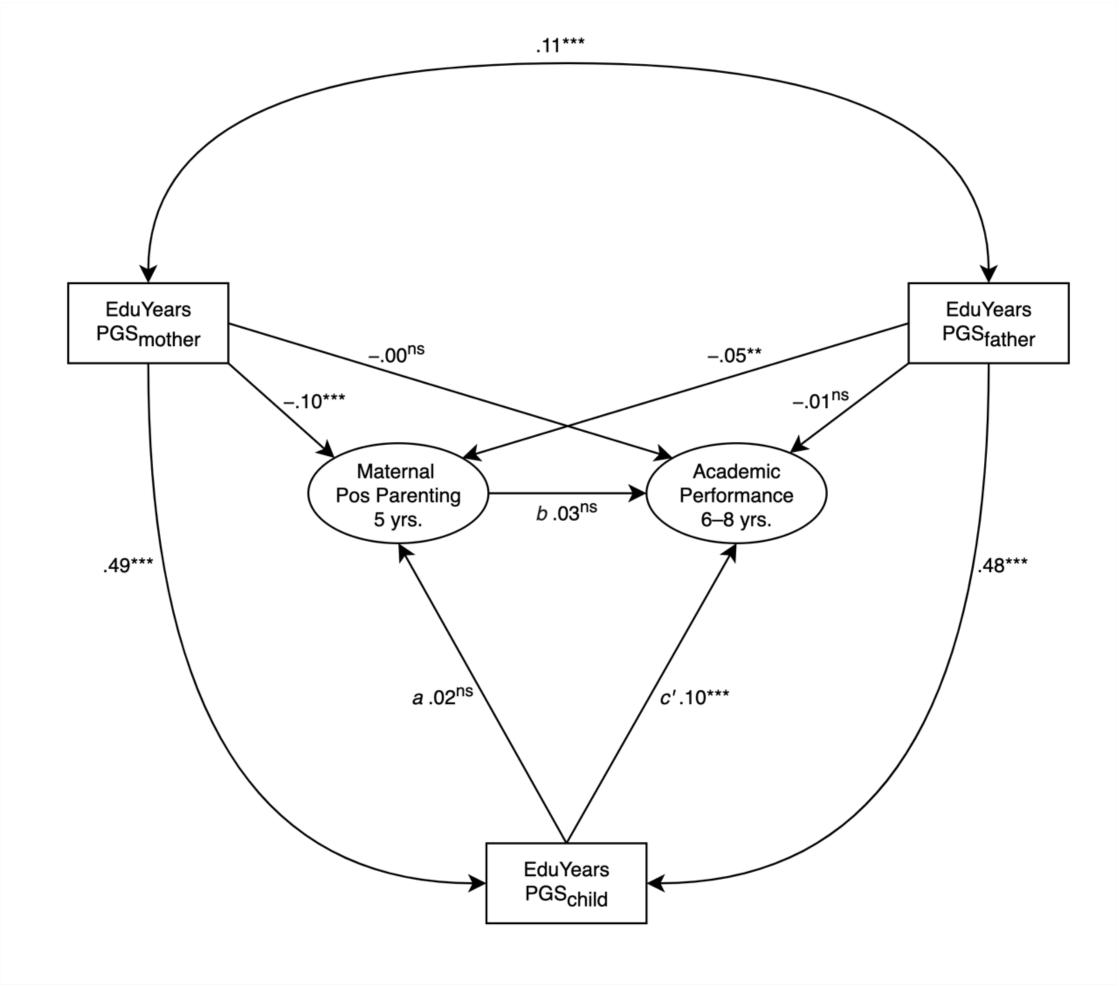
Test of Hypothesis 2: Effect of Child Polygenic Score on Academic Performance Mediated via Maternal Parenting.

Positive Maternal Parenting. To test hypothesis 2, I ran the SEM in Figure 5.2, combining the trio models from Figures 1a and 1c. There was no evidence that maternal positive parenting when children were 5 years old mediated the effects of children’s education-associated genes on their academic performance at 6–8 years. In the model displayed in Figure 5.2, the mediated effect of children’s polygenic scores on their educational performance at 6–8 years old via maternal positive parenting was not statistically significant (*a*b*: $\beta = .00$, 95% CI [–.00, .00], $p = .413$). Maternal positive parenting at 5 years old did not predict children's academic performance at 6–8 years old ($\beta = .03$, 95% CI [–.01, .06], $p = .113$). The model in Figure 5.2 explained 3% of the variance in academic performance at 6–8 years old.

Cognitively Stimulating Maternal Parenting. As a second test of hypothesis 2, I ran the SEM in Figure 5.3, combining the trio models from Figures 1a and 1d. There was no evidence that maternal cognitively stimulating parenting when children were 5 years old mediated the effects of children's education-associated genes on their academic performance at 6–8 years. In the model displayed in Figure 5.3, the mediated effect of children's polygenic scores on their educational performance at 6–8 years old via maternal cognitively stimulating parenting was not statistically significant (*a*b*: $\beta = .00$, 95% CI [–.00, .01], $p = .098$). However, maternal cognitively stimulating parenting at 5 years did predict children's academic performance at 6–8 years old ($\beta = .16$, 95% CI [.13, .18], $p < .001$). The model in Figure 5.3 explained 5% of the variance in academic performance at 6–8 years old.

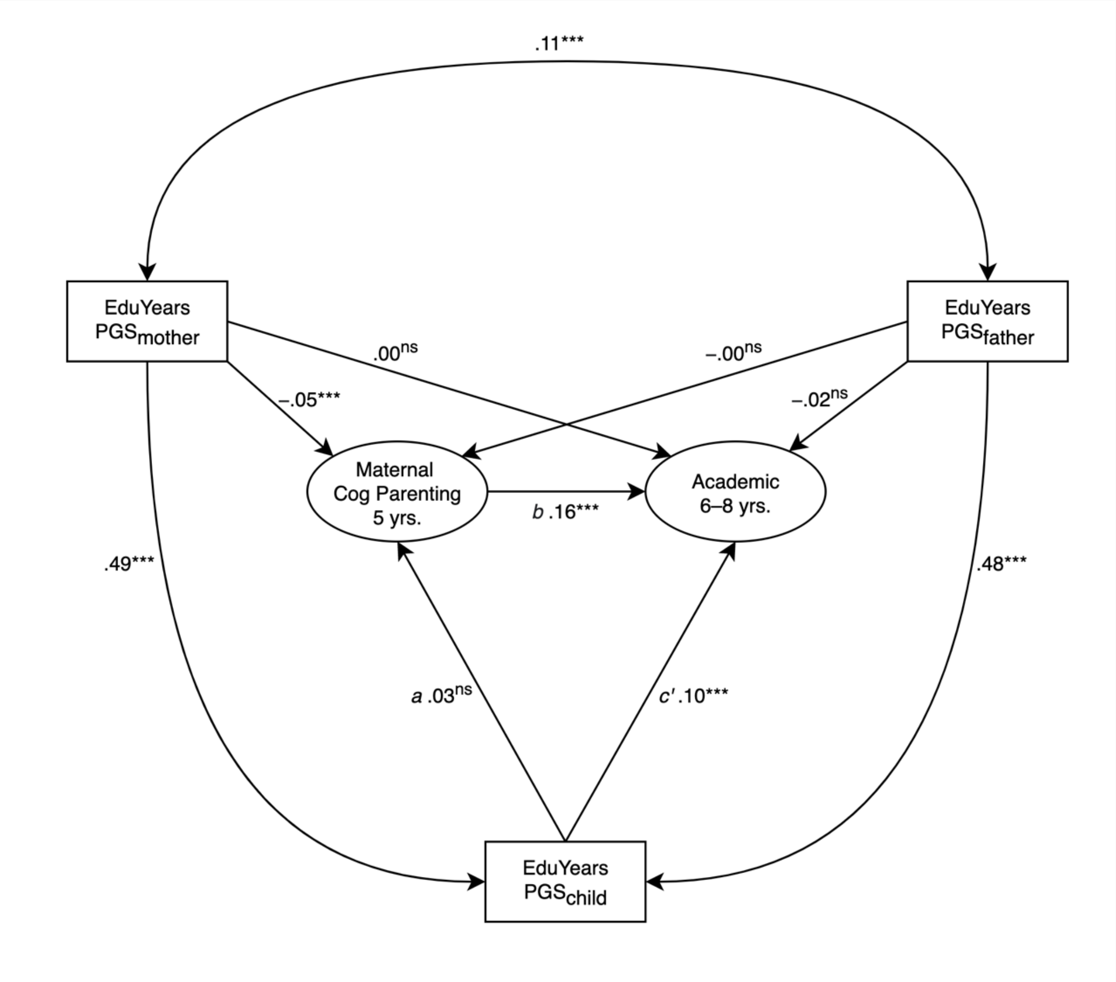
Test of Hypothesis 3: Effect of Child Polygenic Score on Academic Performance Mediated via Language. To test hypothesis 3, I ran the SEMs in Figures 4 and 5, incorporating the trio model from Figure 5.1b into the models displayed in Figures 2 and 3. In the model displayed in Figure 5.4, the mediated effect of children's EduYears polygenic scores on educational performance, via language at 5 years old, was statistically significant in the positive parenting model (*a*b*: $\beta = .01$, 95% CI [.00, .02], $p = .015$), as was the mediated effect via cognitively stimulating parenting, in the model displayed in Figure 5.5, (*a*b*: $\beta = .01$, 95% CI [.00, .01], $p = .016$). As displayed in Figure 5.4, language at 5 years old was positively associated with maternal positive parenting at 5 years old ($\beta = .16$, 95% CI [.13, .18], $p < .001$). As displayed in Figure 5.5, language at 5 years old was positively associated with maternal cognitive stimulation at 5 years old ($\beta = .17$, 95% CI [.15, .19], $p < .001$). The positive parenting model (Figure 5.4) explained 8% of the variance in academic performance at 6–8 years old and the cognitively stimulating parenting model (Figure 5.5) explained 9% of the variance in academic performance.

Figure 5.2 Longitudinal Structural Equation Model Testing Hypothesis 2: Effect of Child Polygenic Score on Academic Performance Mediated via Maternal Positive Parenting



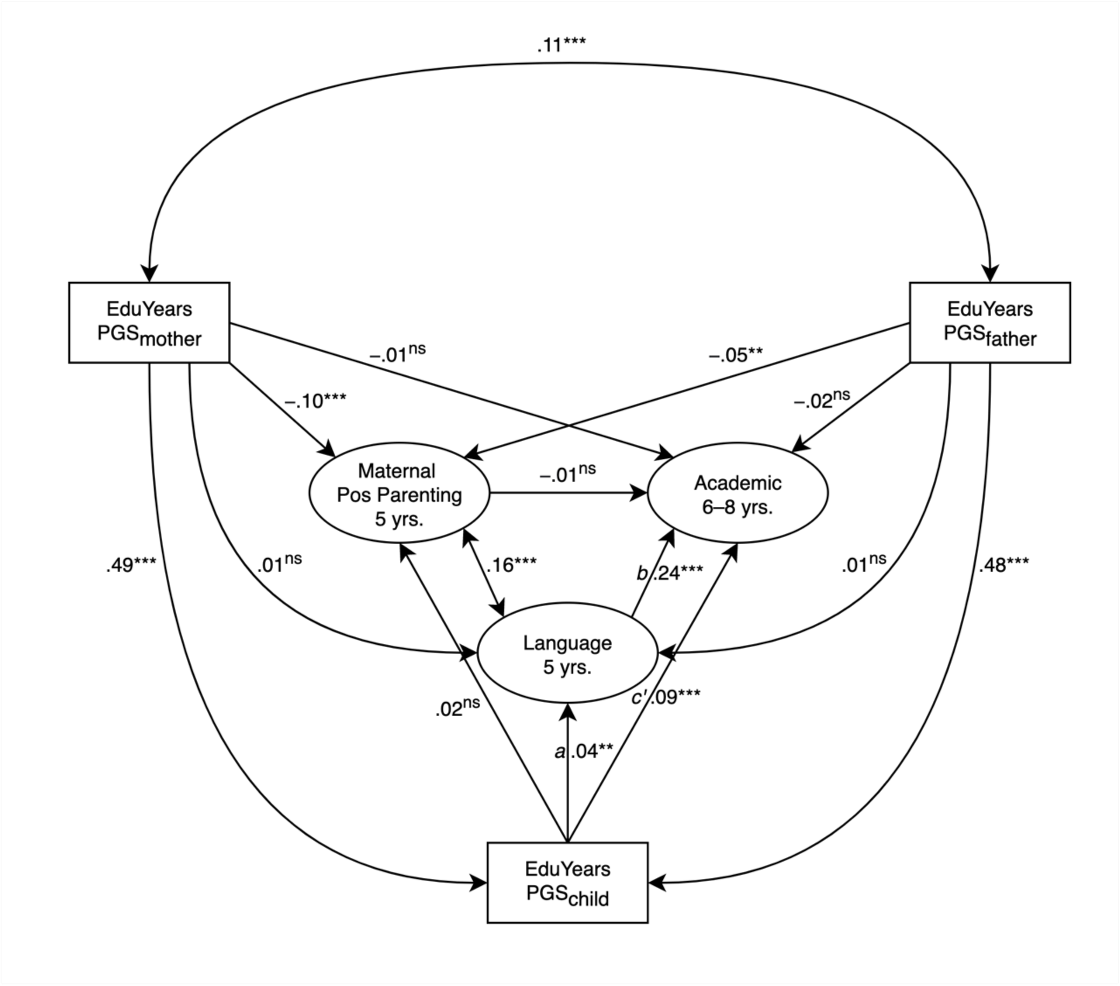
Note: EduYears PGS = years of education polygenic score. Pos = Positive. Model fit: $\chi^2(67) = 975$, $p < .001$, comparative fit index = .98, root mean square error of approximation = .02, standardized root mean square residual = .03, Standardised estimates reported. Child sex and year of birth were included as covariates in the model. $^{ns}p \geq .1$; $^{**}p < .01$; $^{***}p < .001$.

Figure 5.3 Longitudinal Structural Equation Model Testing Hypothesis 2: Effect of Child Polygenic Score on Academic Performance Mediated via Maternal Cognitively Stimulating Parenting



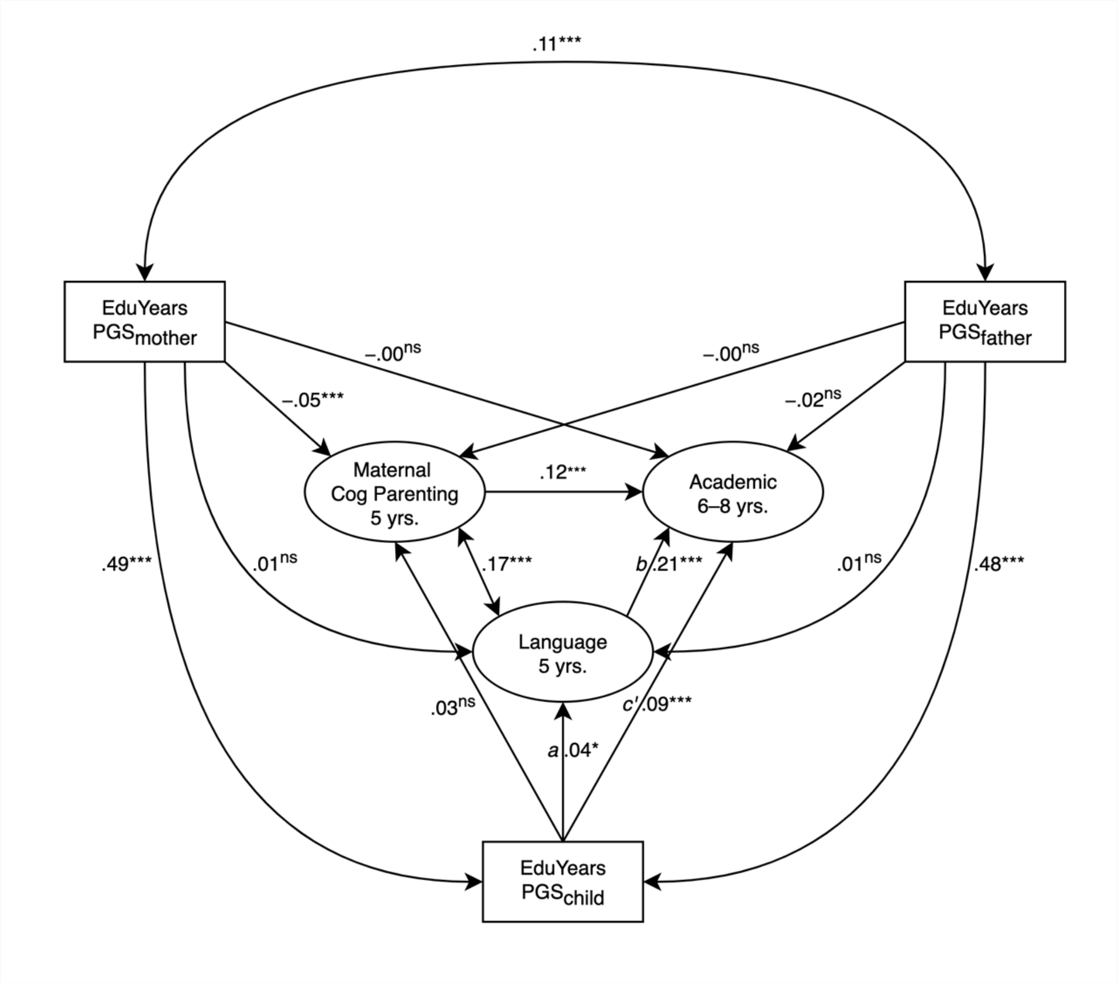
Note: EduYears PGS = years of education polygenic score. Cog = Cognitively stimulating. Model fit: $\chi^2(18) = 242$, $p < .001$, comparative fit index = .99, root mean square error of approximation = .02, standardized root mean square residual = .02, Standardised estimates reported. Child sex and year of birth were included as covariates in the model. $^{ns}p \geq .1$; $^{***}p < .001$.

Figure 5.4 Longitudinal Structural Equation Model Testing Hypothesis 3: Effect of Child Polygenic Score on Academic Performance Mediated via Language



Note: EduYears PGS = years of education polygenic score. Pos = Positive. Model fit: $\chi^2(307) = 20350$, $p < .001$, comparative fit index = .89, root mean square error of approximation = .05, standardized root mean square residual = .03, Standardised estimates reported. Child sex and year of birth were included as covariates in the model. $^{ns}p \geq .1$; $^{**}p < .01$; $^{***}p < .001$.

Figure 5.5 Longitudinal Structural Equation Model Testing Hypothesis 3: Effect of Child Polygenic Score on Academic Performance Mediated via Language



Note: EduYears PGS = years of education polygenic score. Cog = Cognitively stimulating. Model fit: $\chi^2(193) = 19469$, $p < .001$, comparative fit index = .89, root mean square error of approximation = .06, standardized root mean square residual = .04, Standardised estimates reported. Child sex and year of birth were included as covariates in the model. $^{ns}p \geq .1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

5.3.3 Multicollinearity

I tested for multicollinearity by running separate multiple regression models, replicating each of the SEMs that jointly modelled child, mother, and father EduYears polygenic scores using estimated factor scores saved from the SEMs and calculating variance inflation factors (VIF) for child, mother, and father polygenic scores in each regression model. The range in VIF for trio polygenic scores across the regression models was 1.49–2.09, suggesting that multicollinearity did not adversely affect the results.

5.4 Discussion

Drawing on genetic data from the Norwegian Mother, Father, and Child Cohort Study (MoBa), a prospective, longitudinal birth cohort, I examined the unmediated effects of education-associated common genetic variants on children's academic outcomes, as well as mediated effects via early maternal parenting and child language ability. There was no evidence to support the first hypothesis, that there would be evocative effects of children's education-associated genes on maternal parenting. Nor was there any evidence to support the second hypothesis, that the effects of children's education-associated genes on their academic performance would be mediated via parenting. However, there was evidence to support the third hypothesis, that the effects of children's education-associated genes on their academic performance would be mediated via their earlier language performance. Additionally, there was evidence that maternal and paternal polygenic scores directly predicted maternal parenting, in the unexpected (negative) direction, that maternal cognitively stimulating parenting positively predicted children's academic performance, and that language was associated with both positive and cognitively stimulating maternal parenting.

5.4.1 Maternal Parenting

There was no evidence to support my first hypothesis, that there would be evocative effects of children's education-associated common genetic variants on maternal parenting. This is contrary to evidence of evocative effects on parenting of genetic influences underlying educational attainment from research using different behavioural genetics methods, for example multivariate twin modelling (Tucker-Drob & Harden, 2012), the adoption design (Chapter 4 of this thesis), and polygenic scores from mother-child dyads (Wertz et al., 2020). Nor was there evidence to support the second hypothesis. Although children's years of education (EduYears) polygenic scores predicted their academic test performance when they were 6–8 years old, the mediated effect of children's EduYears polygenic scores on academic test performance, via parenting when they were 5 years old, was not statistically significant.

In line with previous research (Wertz et al., 2019; Wertz et al., 2020), mothers' education-associated common genetic variants predicted their parenting. However, contrary to this previous research, which found that children of parents carrying a higher number of education-associated genetic variants were exposed to warmer, more sensitive and stimulating parenting, the present study found that higher maternal education polygenic scores predicted less positive and less cognitively stimulating parenting. It was also the case that fathers' education-associated genes were (unexpectedly) negatively associated with positive maternal parenting (but not cognitively stimulating parenting), suggesting that fathers' genes may evoke differences in their co-parent's positive parenting. The possibility that fathers' genetic differences may influence maternal parenting highlights the general importance of including fathers (and not just mothers) in parenting research, as well as the specific need to incorporate genetic data from both parents when examining evocative effects of children's genes on parenting. As Wertz et al. (2020) highlight, research such as theirs, which examines evocative effects of children's genes on parenting using genetic data from mother-child dyads (without genetic data from fathers), should be interpreted with caution. In such research, the effect of (ostensibly) children's genes on maternal parenting may be confounded by the unmediated effect of fathers' polygenic scores on the maternal parenting of their co-parent.

The unexpected negative associations in the present analyses between parental polygenic scores and positive dimensions of parenting run counter to theories proposed in the wider literature (and empirical evidence supporting them), that higher parental education may encourage positive parenting (such as reading, playing, and modelling positivity and warmth) because it affords parents better access to, and understanding of, information on parenting, child development, and child wellbeing (Davis-Kean, 2005; Davis-Kean et al., 2019). The present findings also diverge with evidence that better educated parents tend to spend more time with their children overall, and engaging them in educationally and developmentally promotive activities (Kalil et al., 2012; Suizzo & Stapleton, 2007), and tend to be more positive and emotionally responsive than those with lower levels of education (Klebanov et al., 1994). Consequently, is unclear whether the unexpected associations in the present work represent a genuine association in the Norwegian population, or a spurious association, induced by, for example, collider bias, which occurs when sample selection invertedly controls for a variable (the ‘collider’) that is independently influenced by both the predictor and outcome, potentially distorting the association between them.

The effects of parents’ polygenic scores on children’s academic test performance were mediated via children’s polygenic scores, suggesting that the bivariate associations between parental polygenic scores and child academic performance are attributable to genetic transmission. However, there was no evidence of genetic nurture effects on children’s academic performance. This is contrary to strong evidence, primarily in older age groups (Wang et al., 2021), but also in one study examining genetic nurture effects on education from 4–7 years old (Armstrong-Carter et al., 2020), of genetic nurture effects on educational outcomes. It is also contrary to results from analyses of the MoBa sample once the children were older (10–13 years old) and using registry data from national exams, rather than parent self-report measures (Isungset et al., 2021).

5.4.2 Language

Children’s EduYears polygenic scores directly predicted language at 5 years old, and there was a statistically significant mediated effect of children’s EduYears polygenic scores on their academic outcomes at 6–8 years old, via language at 5 years old. This is in line with previous research, such as Chapter 3 of this thesis (Austerberry, Fearon, et al., 2022) and Verhoef et al. (2021) and reinforces the view that language may be an early manifestation of genetic effects on academic test performance. Contrary to the lack of association between language and positive parenting in the Early Growth and Development Study (presented in **Chapter 4** of this thesis), language and maternal parenting were positively associated at 5 years old. As the association was cross-sectional, conclusions cannot be drawn about whether this represents a child-to-parent effect or a parent-to-child effect. Previous research suggests that both are plausible in language development (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012). The effects of parental polygenic scores on children’s language were mediated via children’s polygenic scores, suggesting that associations between parent polygenic scores and child language may be explained by genetic transmission. However, there was no evidence of genetic nurture effects on children's language at 5 years old.

5.4.3 Limitations and Future Directions

The findings from this study need to be interpreted in light of several methodological limitations. The first is that polygenic scores, which were used as proxies for participants' genotypes, do not capture all genetic variation that may be relevant to educational attainment. Instead, they represent aggregates of common genetic variants that are weighted for their association with educational attainment in the discovery GWAS (Lee et al., 2018). Consequently, polygenic scores are only able to account for a fraction of the phenotypic variance explained by quantitative genetic studies, such as those using twin and adoption designs. As a result, the reported estimates of genetic effects on parenting, education and language likely underestimate true genetic effects, which may explain why these effect sizes in the present findings were so small. Until more of the "missing heritability" has been accounted for by genetic association studies, it is important to triangulate findings from polygenic scores research with estimates from quantitative genetic studies. The fact that no evocative effects of children's education-associated genes on parenting were detected, even though such effects have been detected in studies using twin and adoption designs, may reflect false negative results in the present study due to the underestimation of genetic effects inherent in these methods. Correspondingly, assortative mating, which was modelled by allowing mother and father polygenic scores to be associated, may have been underestimated. Its underestimation is particularly plausible given that there is robust evidence of positive assortative mating (which occurs when people select mates who are similar to themselves) in educational attainment (Robinson et al., 2017).

The issue of missing heritability also has consequences for the estimation of mediated effects of polygenic scores because of undercorrection for genetic confounding (Pingault et al., 2022). It is a key assumption of mediation models (such as my test of the hypothesis that the association between children's polygenic scores and their educational attainment would be mediated via their early language) that no variables explaining the associations between the three variables (i.e., confounders) are omitted from the model. In single mediator models, if there is an omitted variable that influences both the mediator and the outcome, the path from the mediator to the outcome will be overestimated, in turn exaggerating the mediated effect (Fritz et al., 2016; Judd & Kenny, 1981). Consequently, if the genetic load not captured by the children's EduYears polygenic scores operates as an unmeasured confounder, influencing both children's early language and children's later educational test performance (which seems highly likely), the proportion of the total effect accounted for by the mediated effect will be systematically overestimated.

Finally, missing heritability limits the extent to which passive gene-environment correlation can be controlled for. Although it was a strength of this study that the polygenic scores of both parents were incorporated, the ability to control for genetic confounding was limited to the extent that polygenic scores are not able to capture the full genetic variance. Genetic confounding can only be robustly ruled out by methods such as the Children of Twins design (D'Onofrio et al., 2003) and the adoption design used in Chapter 4 of this thesis (Leve et al., 2019; Leve, Neiderhiser, et al., 2013b). Consequently, what appear to be the effects of the children's polygenic scores may be confounded by the effects of the unmeasured parts of their parents' genomes, not captured by polygenic scores.

The phenotypic measures used were also limited in several respects. First, although genetic data from fathers were included in the analyses, the parenting measures captured maternal parenting only. Consequently, it is unclear whether these results generalize to the parenting of fathers, who are underrepresented overall in developmental research (Phares et al., 2005). Second, the maternal parenting measures were brief self-report items of positive and stimulating parenting. Self-report measures of parenting are vulnerable to rater bias and ceiling effects, which occur when a large proportion of respondents score near the upper limit of a scale so that variance is not measured above a certain level. In the present sample, most mothers rated their parenting as 'often' or 'always' positive, resulting in low variability in the responses. It is unclear whether the results would have remained the same, had it been possible to incorporate detailed observational measures of parenting. Phenotypic measurement is a particular challenge for genomic research as genomic data analysis relies on large sample sizes and it is not usually possible to collect detailed phenotypic measures at the scale required. Language and school performance were based on brief maternal reports by mothers of their children's performance, rather than standardized tests on the child themselves. Future research is needed replicating the present methods using more detailed and objective sources of measurement, such as registry data on grades in national exams.

The MoBa study is Norway-based, and I only included participants in the analyses whose genetic data matched to a European reference panel. I did this because the target data were predominantly European and the EduYears summary statistics used to create the polygenic scores were constructed based on a GWAS of samples of European ancestry. It is a serious limitation of GWAS is that they have been conducted primarily in populations of European descent (Peterson et al., 2019). Polygenic scores currently show poor generalisability in non-European populations and efforts are underway to increase their accuracy across diverse groups (Wang et al., 2022). Once it becomes possible to, the methods from this study should be replicated in different ancestral groups. Currently it remains to be seen whether these results would generalise beyond the Norwegian, European context. Norway is an affluent country with a high standard of living and high equality. For example, Norway had the highest Human Development Index in the world between 1999 (when MoBa recruitment began) and 2016 (when the children in the last wave of recruitment turned 8 years old, which was the latest timepoint I analysed) and has the third highest gross national income per capita (The World Bank, 2023). Compulsory schooling does not begin in Norway until children are 6 years old, however, prior to that most Norwegian children attend kindergarten. In 2004, when children of pregnant mothers recruited to MoBa in 1999 (when recruitment began) reached the age of 5 (the first timepoint used in our analyses), 88% of children in Norway aged 3 and up were in full-time kindergarten, rising to 97% by 2013, when children of mothers recruited in the final year of recruitment (2008) reached 5 years old (The Norwegian Directorate of Education, 2020). It has been hypothesised that the more equal children's educational environments are, the more likely it is that genetic variation will have a greater influence on individual differences in educational phenotypes (Asbury & Plomin, 2013; Scarr-Salapatek, 1971). Consequently, it is plausible that the relatively uniform childcare and schooling environment in Norway—with most children in the analysed age ranges either in full-time kindergarten or schooling—may result in higher heritability estimates for cognitive and educational phenotypes, compared to countries with less equal early caregiving and schooling environments. A further point to note is that the Norwegian mothers and fathers included in the analyses were better educated than the general population in Norway during the study recruitment period (1999–2008). Thus, we cannot assume that the findings from this empirical chapter generalise to lower SES families, particularly as SES appears to moderate genetic effects on intellectual and educational outcomes (Capron & Duyme, 1989; Tucker-Drob & Bates, 2015; Turkheimer et al., 2003) and may causally influence parenting (Akee et al., 2010). Research suggests that participant bias in large-scale volunteer-based studies can distort genetic findings, overestimating SNP effects in educational attainment GWAS (Schoeler et al., 2022). Application of probability weighting can eliminate most of this bias and should be performed in future research (Schoeler et al., 2022).

5.4.4 Conclusion

Contrary to wider research, this study provides no evidence of evocative effects of children's education-associated common genetic variants on maternal parenting. However, it does provide evidence that mothers' and fathers' education-associated genes may influence maternal parenting and that children's education polygenic scores predict their academic performance directly and indirectly via language performance at 5 years old. These findings confirm results from Chapter 2 suggesting that early language ability may be an important mechanism in the pathway from genes to educational outcomes. This offers a guide for promotive and preventative intervention development and research on the causal mechanisms involved in the aetiology of educational performance.

5.5 Acknowledgements

The Norwegian Mother, Father and Child Cohort Study is supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. I am grateful to all the participating families in Norway who take part in this on-going cohort study. I thank the Norwegian Institute of Public Health (NIPH) for generating high-quality genomic data. This research is part of the HARVEST collaboration, supported by the Research Council of

Norway (#229624). For providing genotype data I also thank deCODE Genetics and the NORMENT Centre, funded by the Research Council of Norway (#223273), South East Norway Health Authorities and Stiftelsen Kristian Gerhard Jebsen. I further thank the Center for Diabetes Research, the University of Bergen for providing genotype data and performing quality control and imputation of the data funded by the ERC AdG project SELECTionPREDISPOSED, Stiftelsen Kristian Gerhard Jebsen, Trond Mohn Foundation, the Research Council of Norway, the Novo Nordisk Foundation, the University of Bergen, and the Western Norway Health Authorities.

5.6 References

Armstrong-Carter, E., Trejo, S., Hill, L. J. B., Crossley, K. L., Mason, D., & Domingue, B. W. (2020). The earliest origins of genetic nurture: The prenatal environment mediates the association between maternal genetics and child development. *Psychological Science, 31*(7), 781-791. https://doi.org/10.1177/0956797620917209

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., & Reiss, D. (2022). Early manifestations of intellectual performance: Evidence that genetic effects on later academic test performance are mediated through verbal performance in early childhood. *Child Development, 93*(2), e188-e206. https://doi.org/10.1111/cdev.13706

Austerberry, C., Zayats, T., Ronald, A., Corfield, E. C., Smajlagic, D., Havdahl, A., Andreassen, O. A., Magnus, P., Njølstad, P., & Fearon, P. (2022, February 9). Evocative effects of common genetic variants associated with educational attainment on the early caregiving environment in the Norwegian Mother, Father and Child Cohort Study (MoBa). https://doi.org/10.17605/OSF.IO/84RMN

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20*, 115-140. https://doi.org/10.2307/271084

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*(4498), 1055-1059. https://doi.org/10.1126/science.7195071

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

Corfield, E., Frei, O., Shadrin, A. A., Rahman, Z., Lin, A., Athanasiu, L., Akdeniz, B. C., Hannigan, L., Wootton, R., Austerberry, C., Hughes, A., Tesli, M., Westlye, L. T., Njølstad, P., Magnus, P., Davies, N. M., Appadurai, V., Hemani, G., Hovig, E., Zayats, T., Ask, H., Reichborn-Kjennerud, T., Andreassen, O. A., & Havdahl, A. (2022). The Norwegian Mother, Father, and Child cohort study (MoBa) genotyping data resource: MoBaPsychGen pipeline v.1. Preprint: https://www.biorxiv.org/content/10.1101/2022.06.23.496289v2.

D'Onofrio, B. M., Turkheimer, E. N., Eaves, L. J., Corey, L. A., Berg, K., Solaas, M. H., & Emery, R. E. (2003). The role of the Children of Twins design in elucidating causal relations between parent characteristics and child outcomes. *Journal of Child Psychology and Psychiatry, 44*(8), 1130-1144. https://doi.org/https://doi.org/10.1111/1469-7610.00196

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346. https://doi.org/10.1037/0033-295X.108.2.346

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 8*(3), 430-457. https://doi.org/10.1207/S15328007SEM0803_5

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Psychometric properties of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies, 15*(5), 595-614. https://doi.org/10.1007/s10826-006-9036-y

Evans, D. M., & Davey Smith, G. (2015). Mendelian randomization: New applications in the coming age of hypothesis-free causality. *Annual Review of Genomics and Human Genetics, 16*, 327-350. https://doi.org/10.1146/annurev-genom-090314-050016

Frick, P. J. (1991). *Alabama Parenting Questionnaire*. University of Alabama.

Frick, P. J., Christian, R. E., & Wootton, J. M. (1999). Age trends in the association between parenting practices and conduct problems. *Behavior Modification, 23*(1), 106-128. https://doi.org/10.1177/0145445599231005

Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications, 10*(1), 1776. https://doi.org/10.1038/s41467-019-09718-5

Hadley, P. A., & Rice, M. L. (1993). Parental judgments of preschoolers' speech and language development: A resource for assessment and IEP planning. *Topics in Speech and Language, 14*(04), 278-288. https://doi.org/10.1055/s-2008-1064177

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S. P., Kovas, Y., Corley, R. P., DeFries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*(11), 1112-1120. https://doi.org/10.1038/mp.2009.55

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hummer, R. A., & Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States. *Population bulletin, 68*(1), 1-16.

Isungset, M. A., Conley, D. C., Zachrisson, H. D., Ystrøm, E., Havdahl, A., Njølstad, P. R., & Lyngstad, T. H. (2021). *Social and genetic effects on educational performance in early adolescence*.

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

Kosik, R., Mandell, G., Fan, A. P., Nguyen, T., Chen, J., & Eaton, W. (2018). The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study. *The European Journal of Psychiatry, 32*(2), 53-62. https://doi.org/10.1016/j.ejpsy.2018.01.001

Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*(10), 2048-2056. https://doi.org/10.1177/0956797613486982

Krapohl, E., Hannigan, L. J., Pingault, J.-B., Patel, H., Kadeva, N., Curtis, C., Breen, G., Newhouse, S. J., Eley, T. C., O'Reilly, P. F., & Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences, 114*(44), 11727-11732. https://doi.org/10.1073/pnas.1707178114

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., Yengo, L., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Hinds, D. A., McCreight, J. C., Huber, K. E., Litterman, N. K., McIntyre, M. H., Mountain, J. L., Noblin, E. S., Northover, C. A. M., Pitts, S. J., Sathirapongsasuti, J. F., Sazonova, O. V., Shelton, J. F., Shringarpure, S., Tian, C., Vacic, V., Wilson, C. H., Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., Turley, P., Chen, G.-B., Emilsson, V., Meddens, S. F. W., Oskarsson, S., Pickrell, J. K., Thom, K., Timshel, P., Vlaming, R., Abdellaoui, A., Ahluwalia, T. S., Bacelis, J., Baumbach, C., Bjornsdottir, G., Brandsma, J. H., Concas, M. P., Derringer, J., Furlotte, N. A., Galesloot, T. E., Girotto, G., Gupta, R., Hall, L. M., Harris, S. E., Hofer, E., Horikoshi, M., Huffman, J. E., Kaasik, K., Kalafati, I. P., Karlsson, R., Kong, A., Lahti, J., van der Lee, S. J., Leeuw, C., Lind, P. A., Lindgren, K.-O., Liu, T., Mangino, M., Marten, J., Mihailov, E., Miller, M. B., van der Most, P. J., Oldmeadow, C., Payton, A., Pervjakova, N., Peyrot, W. J., Qian, Y., Raitakari, O., Rueedi, R., Salvi, E., Schmidt, B., Schraut, K. E., Shi, J., Smith, A. V., Poot, R. A., St Pourcain, B., Teumer, A., Thorleifsson, G., Verweij, N., Vuckovic, D., Wellmann, J., Westra, H.-J., Yang, J., Zhao, W., Zhu, Z., Alizadeh, B. Z., Amin, N., Bakshi, A., Baumeister, S. E., Biino, G., Bønnelykke, K., Boyle, P. A., Campbell, H., Cappuccio, F. P., Davies, G., De Neve, J.-E., Deloukas, P., Demuth, I., Ding, J., Eibich, P., Eisele, L., Eklund, N., Evans, D. M., Faul, J. D., Feitosa, M. F., Forstner, A. J., Gandin, I., Gunnarsson, B., Halldórsson,

B. V., Harris, T. B., Heath, A. C., Hocking, L. J., Holliday, E. G., Homuth, G., Horan, M. A., Hottenga, J.-J., de Jager, P. L., Joshi, P. K., Jugessur, A., Kaakinen, M. A., Kähönen, M., Kanoni, S., Keltigangas-Järvinen, L., Kiemeney, L. A. L. M., Kolcic, I., Koskinen, S., Kraja, A. T., Kroh, M., Kutalik, Z., Latvala, A., Launer, L. J., Lebreton, M. P., Levinson, D. F., Lichtenstein, P., Lichtner, P., Liewald, D. C. M., Life Lines Cohort Study, Loukola, A., Madden, P. A., Mägi, R., Mäki-Opas, T., Marioni, R. E., Marques-Vidal, P., Meddens, G. A., McMahon, G., Meisinger, C., Meitinger, T., Milaneschi, Y., Milani, L., Montgomery, G. W., Myhre, R., Nelson, C. P., Nyholt, D. R., Ollier, W. E. R., Palotie, A., Paternoster, L., Pedersen, N. L., Petrovic, K. E., Porteous, D. J., Räikkönen, K., Ring, S. M., Robino, A., Rostapshova, O., Rudan, I., Rustichini, A., Salomaa, V., Sanders, A. R., Sarin, A.-P., Schmidt, H., Scott, R. J., Smith, B. H., Smith, J. A., Staessen, J. A., Steinhagen-Thiessen, E., Strauch, K., Terracciano, A., Tobin, M. D., Ulivi, S., Vaccargiu, S., Quaye, L., van Rooij, F. J. A., Venturini, C., Vinkhuyzen, A. A. E., Völker, U., Völzke, H., Vonk, J. M., Vozzi, D., Waage, J., Ware, E. B., Willemsen, G., Attia, J. R., Bennett, D. A., Berger, K., Bertram, L., Bisgaard, H., Boomsma, D. I., Borecki, I. B., Bültmann, U., Chabris, C. F., Cucca, F., Cusi, D., Deary, I. J., Dedoussis, G. V., van Duijn, C. M., Eriksson, J. G., Franke, B., Franke, L., Gasparini, P., Gejman, P. V., Gieger, C., Grabe, H.-J., Gratten, J., Groenen, P. J. F., Gudnason, V., van der Harst, P., Hayward, C., Hinds, D. A., Hoffmann, W., Hyppönen, E., Iacono, W. G., Jacobsson, B., Järvelin, M.-R., Jöckel, K.-H., Kaprio, J., Kardia, S. L. R., Lehtimäki, T., Lehrer, S. F., Magnusson, P. K. E., Martin, N. G., McGue, M., Metspalu, A., Pendleton, N., Penninx, B. W. J. H., Perola, M., Pirastu, N., Pirastu, M., Polasek, O., Posthuma, D., Power, C., Province, M. A., Samani, N. J., Schlessinger, D., Schmidt, R., Sørensen, T. I. A., Spector, T. D., Stefansson, K., Thorsteinsdottir, U., Thurik, A. R., Timpson, N. J., Tiemeier, H., Tung, J. Y., Uitterlinden, A. G., Vitart, V., Vollenweider, P., Weir, D. R., Wilson, J. F., Wright, A. F., Conley, D. C., Krueger, R. F., Smith, G. D., Hofman, A., Laibson, D. I., Medland, S. E., Meyer, M. N., Yang, J., Johannesson, M., Visscher, P. M., Esko, T., Koellinger, P. D., Cesarini, D., 23andMe Research Team, Cogent, & Social Science

Genetic Association Consortium. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics, 50*(8), 1112-1121. https://doi.org/10.1038/s41588-018-0147-3

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 22*(6), 716-727. https://doi.org/10.1017/thg.2019.66

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J. M., Natsuaki, M. N., & Reiss, D. (2013). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*(1), 412-423. https://doi.org/10.1017/thg.2012.126

Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children’s cognitive development across the first 3 years. *Child Development, 79*(4), 1065-1085. https://doi.org/10.1111/j.1467-8624.2008.01176.x

Magnus, P., Birke, C., Vejrup, K., Haugan, A., Alsaker, E., Daltveit, A. K., Handal, M., Haugen, M., Høiseth, G., Knudsen, G. P., Paltiel, L., Schreuder, P., Tambs, K., Vold, L., & Stoltenberg, C. (2016). Cohort profile update: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology, 45*(2), 382-388. https://doi.org/10.1093/ije/dyw029

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjaerven, R., & Stoltenberg, C. (2006). Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology, 35*(5), 1146-1150. https://doi.org/10.1093/ije/dyl170

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

McVean, G. A., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., Mardis, E. R., Marth, G. T., McVean, G. A., Nickerson, D. A., Schmidt, J. P., Sherry, S. T., Wang, J., Wilson, R. K., Gibbs, R. A., Dinh, H., Kovar, C., Lee, S., Lewis, L., Muzny, D., Reid, J., Wang, M., Wang, J., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Li, G., Li, J., Li, Y., Li, Z., Liu, X., Lu, Y., Ma, X., Su, Z., Tai, S., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Yin, Y., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Zhou, Y., Lander, E. S., Altshuler, D. M., Gabriel, S. B., Gupta, N., Flicek, P., Clarke, L., Leinonen, R., Smith, R. E., Zheng-Bradley, X., Bentley, D. R., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Lehrach, H., Sudbrak, R., Albrecht, M. W., Amstislavskiy, V. S., Borodina, T. A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.-L., Sherry, S. T., McVean, G. A., Mardis, E. R., Wilson, R. K., Fulton, L., Fulton, R., Weinstock, G. M., Durbin, R. M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T. M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Schmidt, J. P., Davies, C. J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Gibbs, R. A., Yu, F., Bainbridge, M., Challis, D., Evani, U. S., Lu, J., Muzny, D., Nagaswamy, U., Reid, J., Sabo, A., Wang, Y., Yu, J., Wang, J., Coin, L. J. M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Y., Li, Z., Lin, H., Liu, B., Luo, R., Qin, N., Shao, H., Wang, B., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Marth, G. T., Garrison, E. P., Kural, D., Lee, W.-P., Fung Leong, W., Ward, A. N., Wu, J., Zhang, M., Lee, C., Griffin, L., Hsieh, C.-H., Mills, R. E., Shi, X., von Grotthuss, M., Zhang, C., Daly, M. J., DePristo, M. A., Altshuler, D. M., Banks, E., Bhatia, G., Carneiro, M. O., del Angel, G., Gabriel, S. B., Genovese, G., Gupta, N., Handsaker, R. E., Hartl, C., Lander, E. S., McCarroll, S. A., Nemesh, J.

C., Poplin, R. E., Schaffner, S. F., Shakir, K., Yoon, S. C., Lihm, J., Makarov, V., Jin, H., Kim, W., Cheol Kim, K., Korbel, J. O., Rausch, T., Flicek, P., Beal, K., Clarke, L., Cunningham, F., Herrero, J., McLaren, W. M., Ritchie, G. R. S., Smith, R. E., Zheng-Bradley, X., Clark, A. G., Gottipati, S., Keinan, A., Rodriguez-Flores, J. L., Sabeti, P. C., Grossman, S. R., Tabrizi, S., Tariyal, R., Cooper, D. N., Ball, E. V., Stenson, P. D., Bentley, D. R., Barnes, B., Bauer, M., Keira Cheetham, R., Cox, T., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J., Shaw, R., Ye, K., Batzer, M. A., Konkel, M. K., Walker, J. A., MacArthur, D. G., Lek, M., Sudbrak, Amstislavskiy, V. S., Herwig, R., Shriver, M. D., Bustamante, C. D., Byrnes, J. K., De La Vega, F. M., Gravel, S., Kenny, E. E., Kidd, J. M., Lacroute, P., Maples, B. K., Moreno-Estrada, A., Zakharia, F., Halperin, E., Baran, Y., Craig, D. W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A. A., Sinari, S. A., Squire, K., Sherry, S. T., Xiao, C., Sebat, J., Bafna, V., Ye, K., Burchard, E. G., Hernandez, R. D., Gignoux, C. R., Haussler, D., Katzman, S. J., James Kent, W., Howie, B., & Ruiz-Linares, A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature, 491*(7422), 56-65. https://doi.org/10.1038/nature11632

Paltiel, L., Anita, H., Skjerden, T., Harbak, K., Bækken, S., Nina Kristin, S., Knudsen, G. P., & Magnus, P. (2014). The biobank of the Norwegian Mother and Child Cohort Study – present status. *Norsk Epidemiologi, 24*(1-2). https://doi.org/10.5324/nje.v24i1-2.1755

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A. C., Kalungi, A., Koen, N., Majara, L., Schwarz, E., Smoller, J. W., Stahl, E. A., Sullivan, P. F., Vassos, E., Mowry, B., Prieto, M. L., Cuellar-Barboza, A., Bigdeli, T. B., Edenberg, H. J., Huang, H., & Duncan, L. E. (2019). Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. *Cell, 179*(3), 589-603. https://doi.org/10.1016/j.cell.2019.08.051

Phares, V., Lopez, E., Fields, S., Kamboukos, D., & Duhig, A. M. (2005). Are fathers involved in pediatric psychology research and treatment? *Journal of Pediatric Psychology, 30*(8), 631-643. https://doi.org/10.1093/jpepsy/jsi050

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.

Prior, M., Bavin, E., & Ong, B. (2011). Predictors of school readiness in five- to six-year-old children from an Australian longitudinal community sample. *Educational Psychology, 31*(1), 3-16. https://doi.org/10.1080/01443410.2010.541048

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*(3), 559-575. https://doi.org/10.1086/519795

R Core Team. (2020). *A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rice, M. L., Wilcox, K. A., Liebhaber, G. K., & Hadley, P. A. (1989). *The speech and Language Assessment Scale*. University of Kansas, USA.

Robinson, M. R., Kleinman, A., Graff, M., Vinkhuyzen, A. A. E., Couper, D., Miller, M. B., Peyrot, W. J., Abdellaoui, A., Zietsch, B. P., Nolte, I. M., van Vliet-Ostaptchouk, J. V., Snieder, H., Alizadeh, B. Z., Boezen, H. M., Franke, L., van der Harst, P., Navis, G., Rots, M., Snieder, H., Swertz, M., Wolffenbuttel, B. H. R., Wijmenga, C., Abecasis, G. R., Absher, D., Alavere, H., Albrecht, E., Allen, H. L., Almgren, P., Amin, N., Amouyel, P., Anderson, D., Arnold, A. M., Arveiler, D., Aspelund, T., Asselbergs, F.

W., Assimes, T. L., Atalay, M., Attwood, A. P., Atwood, L. D., Bakker, S. J. L., Balkau, B., Balmforth, A. J., Barlassina, C., Barroso, I., Basart, H., Bauer, S., Beckmann, J. S., Beilby, J. P., Bennett, A. J., Ben-Shlomo, Y., Bergman, R. N., Bergmann, S., Berndt, S. I., Biffar, R., Di Blasio, A. M., Boehm, B. O., Boehnke, M., Boeing, H., Boerwinkle, E., Bolton, J. L., Bonnefond, A., Bonnycastle, L. L., Boomsma, D. I., Borecki, I. B., Bornstein, S. R., Bouatia-Naji, N., Boucher, G., Bragg-Gresham, J. L., Brambilla, P., Bruinenberg, M., Buchanan, T. A., Buechler, C., Cadby, G., Campbell, H., Caulfield, M. J., Cavalcanti-Proença, C., Cesana, G., Chanock, S. J., Chasman, D. I., Chen, Y. I., Chines, P. S., Clegg, D. J., Coin, L., Collins, F. S., Connell, J. M., Cookson, W., Cooper, M. N., Croteau-Chonka, D. C., Cupples, L. A., Cusi, D., Day, F. R., Day, I. N. M., Dedoussis, G. V., Dei, M., Deloukas, P., Dermitzakis, E. T., Dimas, A. S., Dimitriou, M., Dixon, A. L., Dörr, M., van Duijn, C. M., Ebrahim, S., Edkins, S., Eiriksdottir, G., Eisinger, K., Eklund, N., Elliott, P., Erbel, R., Erdmann, J., Erdos, M. R., Eriksson, J. G., Esko, T., Estrada, K., Evans, D. M., de Faire, U., Fall, T., Farrall, M., Feitosa, M. F., Ferrario, M. M., Ferreira, T., Ferrières, J., Fischer, K., Fisher, E., Fowkes, G., Fox, C. S., Franke, L., Franks, P. W., Fraser, R. M., Frau, F., Frayling, T., Freimer, N. B., Froguel, P., Fu, M., Gaget, S., Ganna, A., Gejman, P. V., Gentilini, D., Geus, E. J. C., Gieger, C., Gigante, B., Gjesing, A. P., Glazer, N. L., Goddard, M. E., Goel, A., Grallert, H., Gräßler, J., Grönberg, H., Groop, L. C., Groves, C. J., Gudnason, V., Guiducci, C., Gustafsson, S., Gyllensten, U., Hall, A. S., Hall, P., Hallmans, G., Hamsten, A., Hansen, T., Haritunians, T., Harris, T. B., van der Harst, P., Hartikainen, A., Hassanali, N., Hattersley, A. T., Havulinna, A. S., Hayward, C., Heard-Costa, N. L., Heath, A. C., Hebebrand, J., Heid, I. M., den Heijer, M., Hengstenberg, C., Herzig, K., Hicks, A. A., Hingorani, A., Hinney, A., Hirschhorn, J. N., Hofman, A., Holmes, C. C., Homuth, G., Hottenga, J., Hovingh, K. G., Hu, F. B., Hu, Y., Huffman, J. E., Hui, J., Huikuri, H., Humphries, S. E., Hung, J., Hunt, S. E., Hunter, D., Hveem, K., Hyppönen, E., Igl, W., Illig, T., Ingelsson, E., Iribarren, C., Isomaa, B., Jackson, A. U., Jacobs, K. B., James, A. L., Jansson, J., Jarick, I., Jarvelin, M., Jöckel, K. H., Johansson, Å., Johnson, T., Jolley, J., Jørgensen, T., Jousilahti, P., Jula, A., Justice, A. E., Kaakinen, M., Kähönen, M., Kajantie, E., Kanoni, S., Kao, W. H. L., Kaplan, L. M., Kaplan, R. C., Kaprio, J., Kapur, K., Karpe, F., Kathiresan, S., Kee, F., Keinanen-Kiukaanniemi, S. M., Ketkar, S., Kettunen, J., Khaw, K., Kiemeney, L. A., Kilpeläinen, T. O., Kinnunen, L., Kivimaki, M., Kivmaki, M., Van der Klauw, M. M., Kleber, M. E., Knowles, J. W., Koenig, W., Kolcic, I., Kolovou, G., König, I. R., Koskinen, S., Kovacs, P., Kraft, P., Kraja, A. T., Kristiansson, K., Krjutškov, K., Kroemer, H. K., Krohn, J. P., Krzelj, V., Kuh, D., Kulzer, J. R., Kumari, M., Kutalik, Z., Kuulasmaa, K., Kuusisto, J., Kvaloy, K., Laakso, M., Laitinen, J. H., Lakka, T. A., Lamina, C., Langenberg, C., Lantieri, O., Lathrop, G. M., Launer, L. J., Lawlor, D. A., Lawrence, R. W., Leach, I. M., Lecoeur, C., Lee, S. H., Lehtimäki, T., Leitzmann, M. F., Lettre, G., Levinson, D. F., Li, G., Li, S., Liang, L., Lin, D., Lind, L., Lindgren, C. M., Lindström, J., Liu, J., Liuzzi, A., Locke, A. E., Lokki, M., Loley, C., Loos, R. J. F., Lorentzon, M., Luan, J., Luben, R. N., Ludwig, B., Madden, P. A., Mägi, R., Magnusson, P. K. E., Mangino, M., Manunta, P., Marek, D., Marre, M., Martin, N. G., The LifeLines Cohort Study, & Genetic Investigation of Anthropometric Traits Consortium. (2017). Genetic evidence of assortative mating in humans. *Nature Human Behaviour, 1*(1), 0016. https://doi.org/10.1038/s41562-016-0016

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/https://doi.org/10.18637/jss.v048.io2

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424-435. https://doi.org/10.2307/1129703

Shelton, K. K., Frick, P. J., & Wootton, J. M. (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology, 25*(3), 317-329. https://doi.org/10.1207/s15374424jccp2503_8

Tierney, N., Cook, D., McBain, M., & Fay, C. (2021). *naniar: Data structures, summaries, and visualisations for missing data. R package version 0.6.1.* https://CRAN.R-project.org/package=naniar

Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene–environment transactions. *Developmental Science, 15*(2), 250-259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P. S., & St Pourcain, B. (2021). The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry, 62*(6), 728-738. https://doi.org/10.1111/jcpp.13327

Wang, B., Baldwin, J. R., Schoeler, T., Cheesman, R., Barkhuizen, W., Dudbridge, F., Bann, D., Morris, T. T., & Pingault, J. B. (2021). Robust genetic nurture effects on education: A systematic review and meta-analysis based on 38,654 families across 8 cohorts. *American Journal of Human Genetics, 108*(9), 1780-1791. https://doi.org/10.1016/j.ajhg.2021.07.010

Wang, Y., Tsuo, K., Kanai, M., Neale, B. M., & Martin, A. R. (2022). Challenges and opportunities for developing more generalizable polygenic risk scores. *Annual Review of Biomedical Data Science, Advance online publication*. https://doi.org/10.1146/annurev-biodatasci-111721-074830

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

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd, L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745-1761. https://doi.org/10.1111/cdev.13329

Chapter 6: Discussion

This thesis had an overarching aim to triangulate findings from a range of genetically sensitive family-based research designs to better understand the earliest manifestations of genetic and environmental influences on complex traits (particularly, psychological traits, developmental milestones, and academic outcomes) and the environmental pathways through which genetic factors might exert their influences. This final chapter begins by summarising the findings from each chapter. Next it considers the extent to which, collectively, the results addressed the overarching questions the thesis aimed to investigate. Finally, the thesis findings are discussed in light of several key limitations and suggestions are made for future research.

6.1 Key Findings

Converging evidence from a range of genetically sensitive research designs has demonstrated that individual differences in virtually all complex traits can be attributed to a combination of genetic and environmental influences. Consequently, the debate has moved on from the question of whether variance in complex traits is attributable to nature *or* nurture. Instead, contemporary behavioural genetics is concerned with questions about the timing of genetic and environmental influences, the interplay of genes and the environment, and which specific genetic and environmental mechanisms might be implicated in the complex pathways to phenotypic variation. Focusing on psychological traits, developmental milestones, and academic outcomes, in early and middle-childhood, this thesis addressed some of these contemporary questions by synthesizing evidence from three genetically sensitive family-based research designs: the classical twin design, the parent-offspring adoption design, and genomic analysis of parent-offspring trios. The results from this thesis provide newly synthesised information about genetic and environmental influences between birth and two years old. The thesis also provides converging evidence from two different family-based research designs (one quantitative genetic and the other genomic) to suggest that early language performance may be a marker of genetic influences on later academic ability. Finally, the results present mixed evidence as to whether genetic differences underlying educational attainment evoke differences in the early caregiving environment and converge in failing to find any evidence that evoked differences in caregiving have a secondary (mediating) influence on academic performance.

6.2 Summary of Each Empirical Chapter

6.2.1 Study 1: Heritability of Psychological Traits and Developmental Milestones in Infancy

The first empirical study of this thesis (**Chapter 2**) aimed to uncover the sources of variation in complex traits in infancy. It presents a systematic review and meta-analysis of all twin studies on psychological traits and developmental milestones in infants (aged 0–2 years) (Austerberry, Mateen, et al., 2022). In a pooled sample of almost 80,000 infant twins, **Chapter 2** found that most categories of traits were moderately or highly heritable, the highest being psychomotor functions (h^2_{pooled}, 59%). Unlike in older age groups, several categories of traits had substantial shared environment estimates in infancy, the highest being mental functions of language (c^2_{pooled}, 59%). All synthesized categories of traits had moderate or high nonshared environment estimates, the highest of which were emotional functions (e^2_{pooled}, 42%) and family relationships (e^2_{pooled}, 42%).

6.2.2 Study 2: The Earliest Manifestations of Genetic Influences on Educational Outcomes

Intellectual and academic abilities are among the most robust predictors of lifelong health and longevity. Despite these assets being moderately to highly heritable in adulthood, as we saw in **Chapter 2**, language and cognitive functions are among the least heritable traits in infancy (Austerberry, Mateen, et al., 2022). The second empirical study (**Chapter 3**) sought to identify the point in development at which genetic influences on academic attainment begin to manifest. Using an adoption sample, **Chapter 3** found that the genetic effects of birth parent academic test performance in adulthood on adoptee academic test performance at 7 years old were mediated via children’s early language performance at 4.5 years old, but not by early executive functioning (EF) (Austerberry, Fearon, et al., 2022). These findings suggest that early language may be a manifestation of genetic advantage for later academic attainment.

6.2.3 Study 3: Evocative effects on the Early Caregiving Environment of Genetic Influences Underlying the Development of Children’s Academic Abilities

It has been long hypothesized that the increasing heritability of intellectual performance across the lifespan is partially attributable to evocative gene-environment correlation (rGE) (Dickens & Flynn, 2001; Plomin et al., 1977; Scarr & McCartney, 1983). The final two empirical studies tested this hypothesis. **Chapter 4** examined, for the first time using an adoption design (which controls for passive rGE), whether the effects of genetic influences on academic test performance at 7 years old are mediated via evoked differences in the early caregiving environment. Guided by the evidence from **Chapter 3** that early childhood language might be a marker of genetic influences on later academic outcomes, **Chapter 4** also examined whether early language differences evoke differences in caregiving. The results demonstrated that birth parent academic performance (used as a proxy for genetic influences on children's academic outcomes) predicted adopted parent positive parenting and household chaos at 6 and 7 years old, suggesting evocative effects on parenting of genetic influences underlying children's academic performance. However, language did not mediate this association, so there was no evidence that language differences evoked these caregiving differences. Furthermore, the apparently evoked differences in parenting did not have a secondary (i.e., mediating) influence on children's academic test performance at 7 years old. Thus, our findings provide no direct evidence to support the hypothesis that evocative rGE is responsible for the increasing heritability of intellectual performance over time.

6.2.4 Study 4: No Evocative Effects on Parenting of Common Genetic Variants Associated with Educational Attainment

Driven by the overarching aim to triangulate findings from different research designs, the fourth empirical chapter (**Chapter 5**) addressed the same research questions as **Chapter 4**, using a different sample (a Norwegian birth cohort) and different methods (genomic analysis of mother-father-child trios). **Chapter 5** examined evocative *r*GE in educational attainment using years of education (EduYears) polygenic scores from children and both of their biological parents. Contrary to the primary hypothesis of this study, there was no evidence of evocative effects of children's education-associated common genetic variants on maternal positive or cognitively stimulating parenting. Nor was there evidence that parenting mediated the effects of polygenic scores on academic outcomes. In line with **Chapter 3**, children's EduYears polygenic scores predicted their academic performance directly and via early language performance, strengthening the evidence that language may be an early marker of genetic influences on academic achievement.

6.3 Overarching Findings

This section synthesises the findings from the four empirical chapters and considers the extent to which, collectively, they address the overarching aims of this thesis. The thesis used genetically sensitive family-based data to address two main research questions. First, what are the earliest manifestations of genetic and environmental influences on psychological and developmental traits. Second, what are the indirect pathways through which genetic factors might exert their influences on academic outcomes. Here we discuss each question, in turn.

6.3.1 Earliest Manifestations of Genetic and Environmental Influences on Complex Traits

Despite the importance of infancy and early childhood as sensitive periods of rapid postnatal growth and development, the earliest indications of genetic and environmental influences on complex psychological traits were not as well understood, at the outset of this thesis, as they were in older age groups. All four empirical studies of this thesis addressed this gap and expanded knowledge on the earliest expressions of genetic and environmental influences on variation in complex psychological traits and developmental milestones.

Early Genetic Influences. Although twin studies had previously been synthesised in a landmark meta-analysis by Polderman et al. (2015), their analysis combined data from infants with data from older children, calculating pooled estimates for 0–11 year olds. Consequently, the sources of variation in complex traits specifically during infancy were not well understood. The first empirical study of the thesis (**Chapter 2**) addressed this gap, conducting the first systematic review and meta-analysis of all twin studies of psychological traits and developmental milestones in infancy (Austerberry, Mateen, et al., 2022). Drawing on a pooled sample of almost 80,000 twins, it demonstrated that all meta-analysed categories of traits had heritability estimates statistically above 0, apart from sleep and language functions. Heritability estimates were high (> 40%) for psychomotor, attention and emotional functions, family relationships (which encompassed infant attachment and dependence) and complex interpersonal interactions (which included phenotypes such as child behavioural problems). These findings were consistent with wider evidence of the ubiquity of genetic influences across the lifespan for virtually all complex traits (Polderman et al., 2015)—what has come to be known as the ‘first law’ of behaviour genetics (Turkheimer, 2000). The results from **Chapter 2** extend this evidence by demonstrating that, even as early as infancy (0–2 years old), a substantial proportion of the variance in most complex psychological traits and developmental milestones appears to be attributable to genetic differences. The high heritability in infancy of attention functions is consistent with the high heritability in older samples of Attention Deficit Hyperactivity Disorder (ADHD) and ADHD traits (Faraone & Larsson, 2019). The relatively high heritability of the subcategories regulating behaviours within interactions (which included phenotypes such as aggression and inadaptability) and social cues in relationships (which included phenotypes such as social autistic traits) is in line with the very high heritability of autism (Tick et al., 2016). The finding that mental functions of language and basic cognitive functions (which encompassed non-verbal cognitive abilities) were two of the least heritable phenotypic domains in infancy is consistent with longitudinal evidence that verbal and non-verbal cognitive abilities have low heritability estimates early in development that increase with age, resulting in moderate-to-high heritability estimates for cognitive and educational performance by adulthood (Davis et al., 2009a; Plomin et al., 1997; Silventoinen et al., 2020).

Following on from the discovery, in **Chapter 2**, that language and cognition were among the least heritable traits in infancy, **Chapter 3** aimed to uncover the earliest manifestations of genetic influences on intellectual and academic outcomes. Using longitudinal data from a parent-offspring adoption study, **Chapter 3** found that language performance at 4.5 years old almost entirely mediated the effects of birth parent intellectual and academic performance (used as a proxy for children’s own genes associated with intellectual and academic skills) on adopted child academic test performance at 7 years old (Austerberry, Fearon, et al., 2022). These findings suggest that language performance at 4.5 years old may be an early marker of genetic propensity for academic performance. Early EF did not mediate any apparently genetic effects. **Chapter 5** also examined whether language mediated genetic effects on academic performance, using genomic (rather than quantitative genetic) methods. It found that language at 5 years old mediated the effects of children's EduYears polygenic scores on their academic test performance at 6–8 years of age, strengthening the evidence that early language may be a marker of genetic propensity for later academic attainment. As is to be expected with comparisons between adoption study findings and results from polygenic score analyses, effect sizes for the genetic effects were much smaller in the polygenic score analyses than in the results from the adoption study analyses (e.g., the proportion of variance in educational attainment explained by the birth mother, parenting and language model in **Chapter 4** was 51%, whereas, despite incorporating genetic data from both parents, the proportion of variance in educational attainment explained by the trio polygenic score, parenting and language model in **Chapter 5** less than one fifth of the size, 9%). These discrepancies are discussed in more detail, in the limitations section, below.

Findings from neither **Chapter 3** nor **Chapter 5** are able to resolve uncertainty about the extent to which early language is a liability index (i.e., there are shared genetic factors that influence both language and subsequent academic test performance) versus a causal mediator of genetic effects on subsequent academic test performance (i.e., limited language development would block the development of the skills necessary to perform well in academic tests) (Kendler & Neale, 2010). Each would have important but different implications for interventions in childhood. Although both suggest that poor language performance is a risk factor for poor academic test performance, the latter suggests that early intervention targeted at language might offset risk, whereas the former might be an indication in favour of more sustained support. Future research should be aimed at testing these alternatives, through longitudinal examination of academic test performance following interventions directly on early language. The findings from **Chapters 3 and 5** also pave the way for genetically informative research testing the causal effects of early language and later academic performance using methods such as Mendelian randomisation, which capitalises on the random segregation of genetic variants to test for causal relationships between genetically correlated phenotypes (Davey Smith & Hemani, 2014).

Early Environmental Influences. The meta-analysis reported in **Chapter 2**, found that, consistent with wider research (Flom et al., 2018), nonshared environment estimates were significantly above 0 for all phenotypic categories meta-analysed, and were high (> 40%) for emotional and attention functions, family relationships, and basic interpersonal interactions. Contrary to evidence in older age groups of the limited influence of the shared environment on individual differences in complex traits (Polderman et al., 2015), in infancy, shared environment estimates had CIs above 0 for several categories of traits. Shared environment estimates were high (> 40%) for language, sleep, growth maintenance, and basic cognitive functions. This is in line with a broader trend in the literature that shared environmental effects on language and cognition appear to be strongest in early development (Davis et al., 2009a; Plomin et al., 1997). There is no straightforward relationship between population genetic and environmental estimates to 'intervention potential'. However, uncovering evidence that individual differences in these traits appear to be particularly attributable in early life to environmental mechanisms is a helpful source of information to guide obesity prevention efforts and efforts to promote intellectual outcomes, which are among the most robust predictors of health and longevity (Deary et al., 2010).

Chapters 4 and 5 aimed to identify specific environmental mechanisms that might be important for early language and academic performance. Specifically, they both examined the potential importance of the early caregiving environment in the aetiological development of children's academic abilities. **Chapter 4** used the parent-offspring adoption design, which can powerfully test the 'pure' effect of the caregiving environment on adoptees because associations between adoptive parent caregiving and adoptee phenotypes are not confounded by shared genes. In these analyses there were no apparent influences of household chaos or adoptive parent warmth on children's language at 4.5 and 6 years old or their academic performance at 7 years old. However, contrary to most previous literature (Adelantado-Renau et al., 2019), there was an association in the unexpected (positive) direction between screen media use at 4.5 years old and children's language performance at 6 years old. In line with **Chapter 4, Chapter 5** found that positive maternal parenting at 5 years old (which captured similar dimensions of parenting to the measure of warmth used in **Chapter 4**) did not predict children's academic test performance at 6–8 years old. However, **Chapter 5** found that, at 5 years old, maternal cognitively stimulating parenting (which was not examined in **Chapter 4**) did predict academic test performance at 6–8 years old. It is surprising, based on the well-established associations between positive parenting and children's intellectual and educational outcomes (Fan & Chen, 2001; Lugo-Gil & Tamis-LeMonda, 2008; Madigan et al., 2019; Wertz et al., 2020), that positive parenting did not predict children's language or academic performance in **Chapters 4 and 5**. The lack of association in analyses using the adoption design (**Chapter 4**) would suggest that previous findings based on polygenic score analyses, for example Wertz et al. (2020), which are not able to robustly control for genetic confounding, may be attributable to passive rGE. Analogously, the similar methods used in **Chapter 5** may have failed to adequately control for passive rGE. Thus, confidence in the observed association, in **Chapter 5**, between cognitively stimulating parenting and children's educational outcomes would be strengthened it were replicated in research using methods such as the parent-offspring adoption design used in **Chapters 3 and 4** or children of twins design (D'Onofrio et al., 2003; Leve et al., 2019; Leve, Neiderhiser, et al., 2013b).

6.3.2 Indirect Pathways Through Which Genetic Factors Might Exert Their Influence

Using two very different research designs, **Chapters 4 and 5** of the thesis addressed the same questions about whether genetic factors indirectly influence academic outcomes via environmental mechanisms. First, they considered whether genetic influences on children's education evoke parenting differences and whether genetic effects on educational outcomes are mediated via parenting. Second, they examined whether children's language performance meallydiated associations between education-associated genetic factors and academic outcomes.

Chapter 4 tested for evocative effects on parenting of education-associated genetic factors using the adoption design, which robustly controls for passive *r*GE. It found that birth parent academic test performance (used as a proxy for genetic influences on children's academic outcomes) positively predicted adoptive parent warmth and household chaos at 6 and 7 years, suggesting evocative effects. This is in line with evidence from one study (Wertz et al., 2020), which found that children's EduYears polygenic scores positively predicted maternal positive parenting (which included warmth), and household chaos after controlling for mothers' own EduYears polygenic scores. However, birth parent intellectual performance did not predict screen media use. This is at odds with evidence from Krapohl et al. (2017) that, at 16 years of age, EduYears polygenic scores negatively predicted 3.4% of the variance in household television watching. Unlike in **Chapter 4**, in **Chapter 5**, which examined associations between children's EduYears polygenic scores and maternal parenting (while controlling for parents' polygenic scores), there was no evidence of evocative effects on maternal positive or cognitively stimulating parenting. This is contrary to evidence from (Tucker-Drob & Harden, 2012), who found that genetic variation almost entirely mediated an observed positive association between children's cognitive performance at 2 years old and cognitively stimulating parenting at 4 years old. Overall, the thesis findings indicated that, in the EGDS sample, adoptive parents may have adjusted their parenting according to their children's genetic propensity for academic achievement, whereas in the MoBa study there was no evidence of this. The causes of similarities and discrepancies between **Chapters 4 and 5**, and wider research, remain unclear and further research is needed before firmer conclusions can be drawn about evocative effects on parenting of education-associated genes.

Chapters 4 and 5 aimed to test the hypothesis that *r*GE accounts for the rising heritability of intellectual outcomes across the lifespan (Dickens & Flynn, 2001; Plomin et al., 1977; Scarr & McCartney, 1983). They did so by examining whether any observed evocative effects of children's education-associated genes on parenting might have a secondary influence on children's academic outcomes. Neither **Chapter 4** nor **Chapter 5** found support for this hypothesis. The effects of genetic influences underlying educational attainment (captured in **Chapter 4** using a measure of birth parent academic test performance and captured in **Chapter 5** using EduYears polygenic scores) on children's educational outcomes in middle childhood were not mediated via any of the measured dimensions of the caregiving environment in either of these two empirical chapters. It is not possible to directly compare these findings with wider research because, although previous research has examined evocative *r*GE in cognitive and academic outcomes, to my knowledge, none has examined whether parenting is a mediator in the association between education-associated genetic influences and educational attainment. Consequently, further research is needed testing this hypothesis.

As language was identified (in **Chapter 3**) as a mediator in the association between education-associated genetic influences and children's academic performance, this thesis hypothesised that early language might be a child phenotype responsible for evoking differences in the early caregiving environment. **Chapter 4** examined this hypothesis, testing whether any observed evocative effects on parenting were mediated via child language. The results revealed no evidence of this. However, **Chapter 5** found that language and maternal parenting (both positive and cognitively stimulating) were associated, suggesting possible effects of children's language on parenting, or of parenting on children's language. As the association was cross-sectional, it was not possible to draw any conclusions about whether this represented child → parent effects, or parent → child effects. Wider literature points to the plausibility of both (Lugo-Gil & Tamis-LeMonda, 2008; Tucker-Drob & Harden, 2012). Further research is needed, testing this hypothesis using genetically sensitive designs and longitudinal measurement.

Although genetic nurture was not a main focus of this thesis, a finding that emerged from the trio analyses in **Chapter 5**, was that parents' polygenic scores predicted children's academic test performance via children's polygenic scores (i.e. genetic transmission), but not genetic nurture. This finding is contrary to strong evidence, reviewed by Wang et al. (2021), primarily in older age groups, but also in one study examining genetic nurture effects on education from 4–7 years old (Armstrong-Carter et al., 2020), of genetic nurture effects on educational outcomes. These findings are also contrary to results from analyses of the MoBa sample once the children were older (10–13 years old) and using registry data from national exams, rather than parent self-report measures (Isungset et al., 2021).

6.4 Limitations and Future Directions

6.4.1 Measurement Issues

Genetic Measurement. Neither of the two quantitative genetic methods (twin and adoption designs) used in **Chapters 2, 3, and 4** involved any direct measurement of the genome. Instead, they relied on relating patterns of phenotypic resemblance to known familial genetic relationships. How closely phenotypic resemblance appeared to mirror patterns of genetic relatedness was used to indirectly estimate genetic influences on the phenotypes in question. Thus, the results revealed nothing about the underlying biological mechanisms involved in associations between putatively genetic factors and complex phenotypes. The polygenic scoring methods used in **Chapter 5** estimate genetic load using genomic data (weighted according to its associations with phenotypes from large discovery GWAS) and thus provide a more direct form of genetic measurement than twin and adoption studies. However, use of polygenic scores does not entirely circumvent issues of indirect measurement. Findings from both quantitative genetic and genomic methods describe the net result of most likely an exceptionally large number of complex gene-environment processes unfolding at multiple levels of biological and social organisation over the course of development. Finding evidence of heritability using any of the methods implemented in this thesis does not imply simple, unmediated, genetic influence on a trait. Many genetic effects likely involve substantial environmental mediation and moderation through mechanisms of gene-environment interaction and correlation (Rutter, 2000). Although this does not undermine the importance for prediction purposes of, for example, identifying associations between polygenic scores and phenotypes, it highlights the gap between indications of genetic association and mechanistic understanding of development. It also highlights the importance of continued research into gene-environment interaction and correlation, akin to the research reported in **Chapters 4 and 5** of the present thesis but involving a wider range of potentially important environmental mechanisms and testing their impact on a variety of phenotypes. Thus far, uncovering evidence of GxE in human populations has proved difficult. Although there may be several reasons for this, a key issue is that GxE effects are inherently dependent on levels of environmental exposure, which may be highly variable from one population to another. GxE effects are also highly dependent on the scaling of the measurements and especially on range limitations, such as floor and ceiling effects (Molenaar & Dolan, 2014). Sophisticated psychometric techniques are being developed to try to address these difficulties (Molenaar & Dolan, 2014), but have not yet been widely adopted. Genomic analyses are also susceptible to bias due to population stratification, which occurs when subgroups within a population differ both genetically and phenotypically, leading to artefactual allele-phenotype associations, which are entirely explained by differences in ancestry. Although attempts were made to statistically control for population stratification by incorporating principal components in the analyses in **Chapter 5**, it is becoming clear that quite subtle ancestral differences may be common and can bias GWAS estimates (Byrne et al., 2020). This, and other forms of bias, continue to be important and active areas of methodological development within the field (Morris et al., 2019).

Although genomic methods offer a more direct form of genetic measurement than quantitative genetic methods (such as twin and adoption studies), the former have not entirely superseded the latter. One of the main reasons for this is that quantitative genetic methods currently explain a far greater proportion of the phenotypic variance than SNP h^2 and polygenic scores constructed from GWAS summary statistics. The results from this thesis were no exception—the variance explained in the twin and adoption study findings reported in **Chapters 2, 3 and 4** was much higher than the variance explained in the polygenic score analyses in **Chapter 5**. As the latter only captured the additive effects of common SNPs weighted for their associations with educational attainment from GWAS, **Chapter 5** likely underestimated the genetic effects being examined (including the effects of assortative mating and genetic confounding). Genetic confounding can only be robustly ruled out by methods such as the adoption design, used in **Chapters 3 and 4** of this thesis, and the children of twins design (D'Onofrio et al., 2003). As a result, the effects in **Chapter 5**, of children's polygenic scores on parenting, language, and educational performance, might have been confounded by the effects of the unmeasured parts of the parental genome, not captured by the parental polygenic scores. Consequently, the effects of the children's polygenic scores were likely simultaneously overestimated, because they did not adequately control for genetic confounding, and underestimated, due to polygenic scores systematically underestimating genetic effects. An implication of the systematic underestimation of genetic influences in **Chapter 5** is that its results may be better interpreted as evidence of likely mechanisms than accurate effect size estimates. Hopefully, with improved methods, such as whole-genome sequencing and increasing sample sizes, GWAS may soon account for a larger fraction of the heritability of behavioural phenotypes. Multivariate analyses incorporating multiple polygenic scores (Plomin & von Stumm, 2018) and the study of the gene-gene and gene-environment interplay are also likely to help explain more of the overall phenotypic variation. However, until the heritability gap is reduced, it remains necessary to take the approach implemented in the present thesis and compare findings from polygenic score research with estimates from family-based quantitative genetic studies, which despite not being based on directly measured DNA, likely better quantify heritability. A caveat is that quantitative genetic studies are capable of systematically overestimating (or underestimating) heritability, if their assumptions are unmet (see the General Introduction of this thesis [**Chapter 1**]). It is also important to hold in mind that, in the present thesis, comparability of findings from the different studies is limited by the fact that each of the different genetically informative family-based methods were applied in a different population. The parent-offspring adoption study in **Chapters 3 and 4** was conducted in the US, whereas the polygenic score analyses in **Chapter 5** were based on data from a Norwegian birth cohort. This issue will be explored in greater detail in the 'Generalisability and Comparability of Findings' section, below.

Phenotypic Measurement. The phenotypic measures used in the thesis were also limited in several respects. The most notable is that there was an overreliance on parent ratings. Almost twice as many independent twin cohorts included in the meta-analysis reported in **Chapter 2** contained parent-reported data ($k_{\text{cohort}} = 22$) than observer-rated data ($k_{\text{cohort}} = 12$). Furthermore, in the three phenotypic categories that contained enough separate data from parents and observers to meta-analyse them independently, heritability estimates were higher for parent ratings. This suggests possible contrast or assimilation effects due to parents exaggerating DZ twin differences and MZ twin similarities (Neale & Stevenson, 1989). Consistent with wider research, nonshared estimates were higher for observer ratings than parent reports, potentially reflecting the importance of each twin's unique experiences in the expression of phenotypes specifically when rated by observers. Alternatively, given that nonshared estimates also include measurement error, higher nonshared estimates for observer-rated phenotypes may reflect increased error in observational measurement. **Chapters 4 and 5** relied on brief parent self-reports of their own parenting, which are vulnerable to rater bias and ceiling effects (which occur when a large proportion of respondents score near the upper limit of a scale so that variance is not measured above a certain level). Indeed, there was evidence of possible ceiling effects in both chapters, as parent self-reports of their parenting were skewed towards positive parenting. It is unclear whether the results from **Chapters 4 and 5** would have been different, had it been possible to incorporate direct observational measures of parenting. Going forward, genetically informative research is needed incorporating observational measures of parenting. Phenotypic measurement is a particular challenge for genomic research as genetic data analysis requires such large sample sizes. Consequently, it is not usually possible to collect detailed phenotypic measures at the scale required. The EGDS adoption study has collected videotaped parent-child observations at the ages examined in this thesis. However, to date, these observations have not been coded. Once these data have been coded, the analyses in **Chapter 4** should be replicated on measures of observed parenting. In **Chapter 5**, children's language and school performance were based on brief maternal reports, rather than direct tests of children themselves. As the ability of genetic measures to predict phenotypes depends partly on reliable phenotypic measurement, the (arguably) better measurement of language and academic performance in EGDS than MoBa may partially explain (along with aforementioned issues linked to missing heritability) why the variance in language and educational performance explained by the polygenic scores in **Chapter 5** was so small compared to the variance explained by birth parent intellectual performance in **Chapters 3 and 4.** Future research is needed replicating the methods from **Chapter 5,** using more detailed and objective sources of measurement, such as registry data on grades in national exams.

Chapter 2 also uncovered, in its systematic review of the infant twin literature, phenotypes that have not been widely measured in infant twin research. Specifically, the most under researched phenotypes, which there was too little data on to meta-analyse, included nonsocial autistic traits and dysregulation, eating behaviours, memory functions, higher-level cognitive functions, and neurobiology. Important new research addressing these gaps is currently underway, such as the BabyTwins study in Sweden, assessing infant brain activity using electroencephalography (Falck-Ytter et al., 2021). However, far more research is needed on the relative contributions of genes and environments to variation in these phenotypes in infancy.

6.4.2 Generalisability and Comparability of Findings

All behaviour genetics findings represent “what is” in a particular sample and cultural context rather than what “could be” in a different context (Plomin et al., 2016). Consequently, we cannot assume that the results from any of the empirical chapters of the thesis provide definitive estimates of genetic and environmental contributions to variation in the phenotypes being examined. As well as curtailing the generalisability of findings, this also limits comparability between empirical chapters. It is quite possible that there are features of the cultural milieu experienced by the US-based adoptees in the EGDS sample in **Chapters 3 and 4** that “transmit” genetic differences into differences in academic test performance to a greater or lesser degree than the cultural context experienced by children participating in the Norway-based MoBa study in **Chapter 5**. Population differences also likely explain why the between-study heterogeneity was so high in the twin study meta-analysis in **Chapter 2**. Crucially, the findings from this thesis were based almost entirely on Western, educated, industrialised, rich, democratic populations. The lack of diversity with regards to ethnicity or ancestry is a serious limitation. In the meta-analysis in **Chapter 2**, 84% of twin pairs were participants in twin studies based in Europe, North America, or Oceania, compared to 16% in Africa, Asia, or South America. The EGDS adoption study, analysed in **Chapters 3 and 4**, is a US-based study and the participating adoptees, birth parents, and adopted parents are predominantly non-Latinx White. The MoBa study, analysed in **Chapter 5**, is based in Norway, and participants were only included in the main analyses if their genetic data had been matched to a European reference panel. This restriction was made because the MoBa participants were predominantly European and the EduYears summary statistics used to construct the polygenic scores were based on a GWAS of samples of European ancestry. Perhaps the biggest limitation of GWAS (upon which polygenic scores rely) is that they have been conducted primarily in populations of European descent (Peterson et al., 2019; Popejoy & Fullerton, 2016). It is a major limitation of psychology research, in general, that it is largely based on samples from wealthy, Western countries, despite individuals from these populations making up for less than 12% of the world's population (Henrich et al., 2010). It is particularly problematic as research suggests that members of such societies, including young children, are not particularly representative when it comes to generalising about human psychology (Henrich et al., 2010). Consequently, it remains unclear whether the results from any of the empirical studies presented in this thesis would generalise beyond the populations they were drawn from. There is evidence from the US and UK that trajectories of language and academic development are not the same for different ethnic groups, and that this may be explained by differences in psychosocial, family and home environments (Saccuzzo et al., 1992; Zilanawala et al., 2016). This highlights a clear need for our methods to be replicated in a more diverse range of populations. This will, in part, require advancements in GWAS research. Some ancestral groups are yet to be sequenced and thus optimal GWAS and genotype imputation for these populations cannot yet be conducted (Tam et al., 2019). Polygenic scores currently show poor generalisability in non-European populations (Wang et al., 2022). For example, the predictive accuracy of the EduYears polygenic score used in **Chapter 5** of this thesis was 11.4% for an ancestrally European sample and 1.6% for an ancestrally African sample (Lee et al., 2018). Poor generalisability across populations can be attributed to population differences in LD (because LD patterns reflect different demographic histories), population specific variation, changes in allele frequency (due to genetic drift and local selection), differences in consanguineous mating, local adaptation, epistasis (gene x gene interaction) due to differences in genetic backgrounds, and GxE interactions (Sirugo et al., 2019). Going forward, hopefully behavioural genetics methods will become more widespread in a diverse range of populations. There are already important efforts underway to increase the accuracy of polygenic scores across diverse groups, for example through data generation and methodological developments in multiancestry polygenic scoring methods (Wang et al., 2022). There are also important twin studies currently taking place in populations outside Europe and North America, for example, the Nigerian

Twin and Sibling registry (Hur, Jeong, et al., 2019), two Japanese twin studies at the Keio Twin Research Center (Ando et al., 2019), the South Korean Twin Registry (Hur, Kang, et al., 2019), and Chinese National Twin Registry (Gao et al., 2019). However, far more research of this kind is needed before conclusions can be drawn about the early manifestations of genetic and environmental influences across the globe or gene-environment correlation in diverse populations.

It also remains to be seen whether the results from this thesis generalise to populations with lower socioeconomic status (SES) than the populations examined in the empirical chapters. Although birth parents in the EGDS sample used in **Chapters 3 and 4** mostly had less than a college education and median household annual incomes below $25,000, the adoptive families raising the adoptees were relatively high SES (typically college educated and with a median household annual income above $100,000). Consequently, the caregiving environments experienced by EGDS adoptees were provided by high SES families. In **Chapter 5**, the Norwegian mothers and fathers included in the analyses were better educated than the general population of women and men in Norway during the study recruitment period (1999–2008). For example, 67% of mothers included in the main analyses, and 52% of included fathers, had completed tertiary education, compared to 21% of women and 21% of men in 1999 (Statistisk Sentralbyrå, 2001), and 29% of women and 25% of men in 2008 (Statistisk Sentralbyrå, 2009). Thus, we cannot assume that the findings from the empirical chapters of the thesis would replicate in lower SES families, particularly as SES appears to moderate genetic effects on intellectual and educational outcomes (Capron & Duyme, 1989; Tucker-Drob & Bates, 2015; Turkheimer et al., 2003) and may causally influence parenting (Akee et al., 2010). Research suggests that adoptees in the US and UK perform better than expected academically (based on their preadoption intelligence scores, education polygenic scores or comparisons with their biological relatives) after being adopted (Cheesman et al., 2020; Duyme et al., 1999; Kendler et al., 2015), indicating mediation or moderation of genetic effects on academic performance by differences in caregiving environments (or the wider social conditions that are associated with them). Consequently, replication of the analyses conducted in this thesis in lower SES samples is needed. Important efforts are underway to study genetic and environmental influences in lower income samples. For example, the Born in Bradford birth cohort in the UK has collected genetic data from children and their parents alongside phenotypic measures of children's health, language and educational outcomes, in a largely South Asian and low SES sample (Bird et al., 2019). It would expand knowledge on the generalisability to lower income samples of the findings from **Chapters 3 and 5** if the mediation of genetic effects on education via language was replicated in this sample.

6.5 Conclusion

This thesis combined evidence from three genetically sensitive family-based research designs: the classical twin design, the parent-offspring adoption design, and genomic analysis of parent-offspring trios. Focusing on psychological, developmental and academic phenotypes in early and middle-childhood, it addressed key contemporary questions about the timing of genetic and environmental influences, the interplay of genes and the environment, and which environmental mechanisms might be implicated in the complex pathways to phenotypic variation. The results provided new pooled estimates of heritability and shared and nonshared environment between birth and 2 years of age. The findings also provide converging evidence from two different family-based research designs (the parent-offspring adoption design and analysis of polygenic scores from parent-offspring trios) to suggest that early language performance may be a marker of genetic influences on later academic ability. Finally, the results present mixed evidence as to whether genetic differences underlying educational attainment evoke differences in the early caregiving environment and converge in failing to find any evidence that evoked differences in early caregiving mediate genetic effects on middle childhood academic performance. In providing this new evidence, the thesis demonstrates that genetically sensitive family-based research can be leveraged not only to further knowledge of genetic sources of variation in complex trait development but also to better understand environmental pathways to complex trait variation.

6.6 References

Adelantado-Renau, M., Moliner-Urdiales, D., Cavero-Redondo, I., Beltran-Valls, M. R., Martínez-Vizcaíno, V., & Álvarez-Bueno, C. (2019). Association between screen media use and academic performance among children and adolescents: A systematic review and meta-analysis. *JAMA Pediatrics, 173*(11), 1058-1067. https://doi.org/10.1001/jamapediatrics.2019.3176

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

Ando, J., Fujisawa, K. K., Hiraishi, K., Shikishima, C., Kawamoto, T., Nozaki, M., Yamagata, S., Takahashi, Y., Suzuki, K., Someya, Y., Ozaki, K., Deno, M., Tanaka, M., Sasaki, S., Toda, T., Kobayashi, K., Sakagami, M., Okada, M., Kijima, N., Takizawa, R., & Murayama, K. (2019). Psychosocial twin cohort studies in Japan: The Keio Twin Research Center (KoTReC). *Twin Research and Human Genetics, 22*(6), 591-596. https://doi.org/10.1017/thg.2019.109

Bird, P. K., McEachan, R. R. C., Mon-Williams, M., Small, N., West, J., Whincup, P., Wright, J., Andrews, E., Barber, S. E., Hill, L. J. B., Lennon, L., Mason, D., Shire, K. A., Waiblinger, D., Waterman, A. H., Lawlor, D. A., & Pickett, K. E. (2019). Growing up in Bradford: Protocol for the age 7–11 follow up of the Born in Bradford birth cohort. *BMC Public Health, 19*(1), 939. https://doi.org/10.1186/s12889-019-7222-2

Byrne, R. P., van Rheenen, W., van den Berg, L. H., Veldink, J. H., McLaughlin, R. L., & Project MinE ALS Gwas Consortium. (2020). Dutch population structure across space, time and GWAS design. *Nature Communications, 11*(1), 4556. https://doi.org/10.1038/s41467-020-18418-4

Capron, C., & Duyme, M. (1989). How much can we raise or lower I.Q.? The joint effects genetic and prenatal factors versus postnatal factors. *Journal de Genetic Human, 37*(2), 115-118.

Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., Eley, T. C., & Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological Science, 31*(5), 582-591. https://doi.org/10.1177/0956797620904450

D'Onofrio, B. M., Turkheimer, E. N., Eaves, L. J., Corey, L. A., Berg, K., Solaas, M. H., & Emery, R. E. (2003). The role of the Children of Twins design in elucidating causal relations between parent characteristics and child outcomes. *Journal of Child Psychology and Psychiatry, 44*(8), 1130-1144. https://doi.org/https://doi.org/10.1111/1469-7610.00196

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

Davis, O. S., Haworth, C. M., & Plomin, R. (2009). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science, 20*(10), 1301-1308. https://doi.org/10.1111/j.1467-9280.2009.02433.x

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346. https://doi.org/10.1037/0033-295X.108.2.346

Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of “dull children”?: A late adoption study. *Proceedings of the National*

Academy of Sciences, 96(15), 8790-8794. https://doi.org/10.1073/pnas.96.15.8790

Falck-Ytter, T., Hamrefors, L., Siqueiros Sanches, M., Portugal, A. M., Taylor, M., Li, D., Viktorsson, C., Hardiansyah, I., Myers, L., Westberg, L., Bölte, S., Tammimies, K., & Ronald, A. (2021). The Babytwins Study Sweden (BATSS): A multi-method infant twin study of genetic and environmental factors influencing infant brain and behavioral development. *Twin Research and Human Genetics, 24*(4), 217-227. https://doi.org/10.1017/thg.2021.34

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1-22. https://doi.org/10.1023/A:1009048817385

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

Flom, M., Wang, M., Uccello, K. J., & Saudino, K. J. (2018). Parent- and observer-rated positive affect in early childhood: Genetic overlap and environmental specificity. *Behavior Genetics, 48*(6), 432-439. https://doi.org/10.1007/s10519-018-9924-0

Gao, W., Cao, W., Lv, J., Yu, C., Wu, T., Wang, S., Meng, L., Wang, D., Wang, Z., Pang, Z., Yu, M., Wang, H., Wu, X., Dong, Z., Wu, F., Jiang, G., Wang, X., Liu, Y., Deng, J., Lu, L., & Li, L. (2019). The Chinese National Twin Registry: A 'gold mine' for scientific research. *Journal of Internal Medicine, 286*(3), 299-308. https://doi.org/https://doi.org/10.1111/joim.12926

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Hur, Y. M., Jeong, H. U., Kang, M. C., Ajose, F., Kim, J. W., Beck, J. J., Hottenga, J. J., Mbarek, H., Finnicum, C. T., Ehli, E. A., Martin, N. G., de Geus, E. J., Boomsma, D. I., Davies, G. E., & Bates, T. (2019). The Nigerian Twin and Sibling Registry: An update. *Twin Research and Human Genetics, 22*(6), 637-640. https://doi.org/10.1017/thg.2019.110

Hur, Y. M., Kang, M. C., Jeong, H.-U., Kang, I. C., & Kim, J. W. (2019). The South Korean Twin Registry. *Twin Research and Human Genetics, 22*(6), 606-608. https://doi.org/10.1017/thg.2019.115

Kendler, K. S., & Neale, M. C. (2010). Endophenotype: A conceptual analysis. *Molecular Psychiatry, 15*(8), 789-797. https://doi.org/10.1038/mp.2010.8

Kendler, K. S., Turkheimer, E., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proceedings of the National Academy of Sciences, 112*(15), 4612-4617. https://doi.org/10.1073/pnas.1417106112

Krapohl, E., Hannigan, L. J., Pingault, J.-B., Patel, H., Kadeva, N., Curtis, C., Breen, G., Newhouse, S. J., Eley, T. C., O'Reilly, P. F., & Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences, 114*(44), 11727-11732. https://doi.org/10.1073/pnas.1707178114

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., Yengo, L., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Hinds, D. A., McCreight, J. C., Huber, K. E., Litterman, N. K., McIntyre, M. H., Mountain, J. L., Noblin, E. S., Northover, C. A. M., Pitts, S. J., Sathirapongsasuti, J. F., Sazonova, O. V., Shelton, J. F., Shringarpure, S., Tian, C., Vacic, V., Wilson, C. H., Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., Turley, P., Chen, G.-B., Emilsson, V., Meddens, S. F. W., Oskarsson, S., Pickrell, J. K., Thom, K., Timshel, P., Vlaming, R., Abdellaoui, A., Ahluwalia, T. S., Bacelis, J., Baumbach, C., Bjornsdottir, G., Brandsma, J. H., Concas, M. P., Derringer, J., Furlotte, N. A., Galesloot, T. E., Girotto, G., Gupta, R., Hall, L. M., Harris, S. E., Hofer, E., Horikoshi, M., Huffman, J. E., Kaasik, K., Kalafati, I. P., Karlsson, R., Kong, A., Lahti, J., van der Lee, S. J., Leeuw, C., Lind, P. A., Lindgren, K.-O., Liu, T., Mangino, M., Marten, J., Mihailov, E., Miller, M. B., van der Most, P. J., Oldmeadow, C., Payton, A., Pervjakova, N., Peyrot, W. J., Qian, Y., Raitakari, O., Rueedi, R., Salvi, E., Schmidt, B., Schraut, K. E., Shi, J., Smith, A. V., Poot, R. A., St Pourcain, B., Teumer, A., Thorleifsson, G., Verweij, N., Vuckovic, D., Wellmann, J., Westra, H.-J., Yang, J., Zhao, W., Zhu, Z., Alizadeh, B. Z., Amin, N., Bakshi, A., Baumeister, S. E., Biino, G., Bønnelykke, K., Boyle, P. A., Campbell, H., Cappuccio, F. P., Davies, G., De Neve, J.-E., Deloukas, P., Demuth, I., Ding, J., Eibich, P., Eisele, L., Eklund, N., Evans, D. M., Faul, J. D., Feitosa, M. F., Forstner, A. J., Gandin, I., Gunnarsson, B., Halldórsson, B. V., Harris, T. B., Heath, A. C., Hocking, L. J., Holliday, E. G., Homuth, G., Horan, M. A., Hottenga, J.-J., de Jager, P. L., Joshi, P. K., Jugessur, A., Kaakinen, M. A., Kähönen, M., Kanoni, S., Keltigangas-Järvinen, L., Kiemeney, L. A. L. M., Kolcic, I., Koskinen, S., Kraja, A. T., Kroh, M., Kutalik, Z., Latvala, A., Launer, L. J., Lebreton, M. P., Levinson, D. F., Lichtenstein, P., Lichtner, P., Liewald, D. C. M., Life Lines Cohort Study, Loukola, A., Madden, P. A., Mägi, R., Mäki-Opas, T., Marioni, R. E., Marques-Vidal, P., Meddens, G. A., McMahon, G., Meisinger, C., Meitinger, T., Milaneschi, Y., Milani, L., Montgomery, G. W., Myhre, R., Nelson, C. P., Nyholt, D. R., Ollier, W. E. R., Palotie, A., Paternoster, L., Pedersen, N. L., Petrovic, K. E., Porteous, D. J., Räikkönen, K., Ring, S. M., Robino, A., Rostapshova, O., Rudan, I., Rustichini, A., Salomaa, V., Sanders, A. R., Sarin, A.-P., Schmidt, H., Scott, R. J., Smith, B. H., Smith, J. A., Staessen, J. A., Steinhagen-Thiessen, E., Strauch, K., Terracciano, A., Tobin, M. D., Ulivi, S., Vaccargiu, S., Quaye, L., van

Rooij, F. J. A., Venturini, C., Vinkhuyzen, A. A. E., Völker, U., Völzke, H., Vonk, J. M., Vozzi, D., Waage, J., Ware, E. B., Willemsen, G., Attia, J. R., Bennett, D. A., Berger, K., Bertram, L., Bisgaard, H., Boomsma, D. I., Borecki, I. B., Bültmann, U., Chabris, C. F., Cucca, F., Cusi, D., Deary, I. J., Dedoussis, G. V., van Duijn, C. M., Eriksson, J. G., Franke, B., Franke, L., Gasparini, P., Gejman, P. V., Gieger, C., Grabe, H.-J., Gratten, J., Groenen, P. J. F., Gudnason, V., van der Harst, P., Hayward, C., Hinds, D. A., Hoffmann, W., Hyppönen, E., Iacono, W. G., Jacobsson, B., Järvelin, M.-R., Jöckel, K.-H., Kaprio, J., Kardia, S. L. R., Lehtimäki, T., Lehrer, S. F., Magnusson, P. K. E., Martin, N. G., McGue, M., Metspalu, A., Pendleton, N., Penninx, B. W. J. H., Perola, M., Pirastu, N., Pirastu, M., Polasek, O., Posthuma, D., Power, C., Province, M. A., Samani, N. J., Schlessinger, D., Schmidt, R., Sørensen, T. I. A., Spector, T. D., Stefansson, K., Thorsteinsdottir, U., Thurik, A. R., Timpson, N. J., Tiemeier, H., Tung, J. Y., Uitterlinden, A. G., Vitart, V., Vollenweider, P., Weir, D. R., Wilson, J. F., Wright, A. F., Conley, D. C., Krueger, R. F., Smith, G. D., Hofman, A., Laibson, D. I., Medland, S. E., Meyer, M. N., Yang, J., Johannesson, M., Visscher, P. M., Esko, T., Koellinger, P. D., Cesarini, D., 23andMe Research Team, Cogent, & Social Science Genetic Association Consortium. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics, 50*(8), 1112-1121. https://doi.org/10.1038/s41588-018-0147-3

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 22*(6), 716-727. https://doi.org/10.1017/thg.2019.66

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J. M., Natsuaki, M. N., & Reiss, D. (2013). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin*

Research and Human Genetics, 16(1), 412-423. https://doi.org/10.1017/thg.2012.126

Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Development, 79*(4), 1065-1085. https://doi.org/10.1111/j.1467-8624.2008.01176.x

Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., & Jenkins, J. M. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics, 144*(4). https://doi.org/10.1542/peds.2018-3556

Molenaar, D., & Dolan, C. V. (2014). Testing systematic genotype by environment interactions using item level data. *Behavior Genetics, 44*(3), 212-231. https://doi.org/10.1007/s10519-014-9647-9

Morris, T. T., Davies, N. M., Hemani, G., & Smith, G. D. (2019). Why are education, socioeconomic position and intelligence genetically correlated? *bioRxiv*, 630426. https://doi.org/10.1101/630426

Neale, M. C., & Stevenson, J. (1989). Rater bias in the EASI temperament scales: A twin study. *Journal of Personality and Social Psychology, 56*(3), 446-455. https://doi.org/10.1037//0022-3514.56.3.446

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A. C., Kalungi, A., Koen, N., Majara, L., Schwarz, E., Smoller, J. W., Stahl, E. A., Sullivan, P. F., Vassos, E., Mowry, B., Prieto, M. L., Cuellar-Barboza, A., Bigdeli, T. B., Edenberg, H. J., Huang, H., & Duncan, L. E. (2019). Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. *Cell, 179*(3), 589-603. https://doi.org/10.1016/j.cell.2019.08.051

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science, 11*(1), 3-23. https://doi.org/10.1177/1745691615617439

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*(2), 309-322.

Plomin, R., Fulker, D. W., Corley, R. P., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 Years: A parent-offspring adoption study. *Psychological Science, 8*(6), 442-447. https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

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

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics, 47*(7), 702-709. https://doi.org/10.1038/ng.3285

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

Rutter, M. (2000). Psychosocial influences: Critiques, findings, and research needs. *Development and Psychopathology, 12*(3), 375-405. https://doi.org/10.1017/S0954579400003072

Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal versus performance IQs for gifted African-American, Caucasian, Filipino, and Hispanic children. *Psychological Assessment, 4*(2), 239. https://doi.org/10.1037/1040-3590.4.2.239

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424-435. https://doi.org/10.2307/1129703

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., Tomizawa, R., Watanabe, M., Sakai, N., Rebato, E., Busjahn, A., Tyler, J., Hopper, J. L., Ordoñana, J. R., Sánchez-Romera, J. F., Colodro-Conde, L., Calais-Ferreira, L., Oliveira, V. C., Ferreira, P. H., Medda, E., Nisticò, L., Toccaceli, V., Derom, C. A., Vlietinck, R. F., Loos, R. J. F., Siribaddana, S. H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Duncan, G. E., Buchwald, D., Tynelius, P., Rasmussen, F., Tan, Q., Zhang, D., Pang, Z., Magnusson, P. K. E., Pedersen, N. L., Dahl Aslan, A. K., Hwang, A. E., Mack, T. M., Krueger, R. F., McGue, M., Pahlen, S., Brandt, I., Nilsen, T. S., Harris, J. R., Martin, N. G., Medland, S. E., Montgomery, G. W., Willemsen, G., Bartels, M., van Beijsterveldt, C. E. M., Franz, C. E., Kremen, W. S., Lyons, M. J., Silberg, J. L., Maes, H. H., Kandler, C., Nelson, T. L., Whitfield, K. E., Corley, R. P., Huibregtse, B. M., Gatz, M., Butler, D. A., Tarnoki, A. D., Tarnoki, D. L., Park, H. A., Lee, J., Lee, S. J., Sung, J., Yokoyama, Y., Sørensen, T. I. A., Boomsma, D. I., & Kaprio, J. (2020). Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. *Scientific Reports, 10*(1), 12681. https://doi.org/10.1038/s41598-020-69526-6

Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell, 177*(1), 26-31. https://doi.org/10.1016/j.cell.2019.02.048

Statistisk Sentralbyrå. (2001). Educational attainment of the population, 1 October 1999. https://www.ssb.no/en/utdanning/statistikker/utniv/arkiv/2001-03-07

Statistisk Sentralbyrå. (2009). *Educational attainment of the population, 1 October 2008*. https://www.ssb.no/en/utdanning/statistikker/utniv/arkiv/2009-08-25

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics, 20*(8), 467-484. https://doi.org/10.1038/s41576-019-0127-1

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

Tucker-Drob, E. M., & Bates, T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138-149. https://doi.org/10.1177/0956797615612727

Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene–environment transactions. *Developmental Science, 15*(2), 250-259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 9*(5), 160-164. https://doi.org/10.1111/1467-8721.00084

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, II. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science, 14*(6), 623-628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

Wang, Y., Tsuo, K., Kanai, M., Neale, B. M., & Martin, A. R. (2022). Challenges and opportunities for developing more generalizable polygenic risk scores. *Annual Review of Biomedical Data Science, Advance online publication*. https://doi.org/10.1146/annurev-biodatasci-111721-074830

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd, L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745-1761. https://doi.org/10.1111/cdev.13329

Zilanawala, A., Kelly, Y., & Sacker, A. (2016). Ethnic differences in longitudinal latent verbal profiles in the millennium cohort study. *European Journal of Public Health, 26*(6), 1011-1016. https://doi.org/10.1093/eurpub/ckw184

Appendix A

A.1 Supplementary Methods

A.1.1 Classical Twin Design

I searched databases for results from classical twin studies. The classical twin design is based on the comparison of phenotypic similarity within monozygotic (MZ) twins and dizygotic (DZ) twins. As MZ twins share 100% of their segregating genome and DZ twins share on average 50%, a higher degree of phenotypic similarity within MZ than DZ twins indicates genetic influence. When DZ twins show within-pair similarity greater than half of that of MZ twins, this suggests shared environmental influences (defined as environmental influences that make children growing up in the same family similar). Differences between MZ twins are attributed to nonshared environmental influences, which are influences of aspects of the environment that children growing up in the same family do not share.

A.1.2 Quality Assessment

Papers were quality assessed using a modified version of the checklist for assessing the quality of quantitative studies (Kmet et al., 2004), which rates studies on a scale from 0 to 2 (2 = Yes; 1 = Partial; 0 = No; NA = Not Applicable) on the following 14 items:

1. Question/objective sufficiently described?
2. Study design evident and appropriate?
3. Method of subject/comparison group selection or source of information/input variables described and appropriate?
4. Subject (and comparison group, if applicable) characteristics sufficiently described?
5. If interventional and random allocation was possible, was it described?
6. If interventional and blinding of investigators was possible, was it reported?
7. If interventional blinding of subjects was possible, was it reported?
8. Outcome and (if applicable) exposure measure(s) well defined and robust to measurement/misclassification bias? Means of assessment reported?
9. Sample size appropriate?
10. Analytic methods described/justified and appropriate?

11. Some estimate of variance is reported for the main results?
12. Controlled for confounding?
13. Results reported in sufficient detail?
14. Conclusion supported by the results?

The modifications made to make the tool suitable for assessing the quality of twin studies were: (1) removal of items 5 and 7, because classical twin studies do not include interventional and random allocation or interventional blinding of subjects; (2) removal of item 12 from the assessment, as the confounds that I was most interested in are controlled for by the twin design itself (e.g. age, genetic influences, shared rearing environment); (3) adaption of item 6 (rewording it to: "If possible, were investigators blind to zygosity or task outcome?"); (4) introduction of an additional item ("Sample included only same-sex twins, or same-sex and opposite-sex twins were separated in the analyses?"), because, if opposite-sex twins are treated more differently than same-sex twins, inclusion of data from opposite-sex twins introduces an additional source of difference between MZ and DZ twins and can lead to inflated heritability estimates.

A.1.3 Duplicate Information

It is common for univariate twin results from a single study to be reported in more than one paper because multivariate twin studies often report univariate estimates for all included phenotypes. Consequently, when one phenotype from a twin sample is analysed in relation to different phenotypes across more than one paper, each paper will typically report univariate estimates for the focal phenotype. As a result, I searched extracted data for overlapping/duplicate information (e.g., papers reporting estimates relating to the same phenotype, measure, age, and cohort) and excluded duplicates from the meta-analysis (prioritizing inclusion of estimates with larger and more recently published estimates).

A.1.4 Three-Level Multilevel Random Effects Meta-Analysis

Before running the meta-analyses, I prepared the extracted data. If studies reported only heritability, shared and nonshared environment (ACE) estimates or both ACE estimates and concordances (not correlations), I used Falconer's formulas to convert the ACE estimates into correlations.(Falconer, 1960) Variances were then calculated using the within-twin correlations (rMZ and rDZ) and sample sizes. The correlations and variances were then used in the three-level multilevel random effects meta-analyses. I conducted three-level multilevel random effects models to facilitate the simultaneous estimation of multiple dependent effect sizes within a single cohort. This allowed me to include all data published in studies that reported estimates from same-sex and opposite-sex twins separately (and sometimes same-sex male and same-sex female, separately) as well as those that reported multiple assessments (e.g., repeated measures) or multiple different measures of the same or similar phenotypes—all of which are common occurrences in large twin studies and registries. Although ideally, I would have incorporated the correlations among dependent effect sizes, this was not possible as this information was not usually reported. Evidence suggest that the hierarchical approach estimates effects well, even without such information (Van den Noortgate et al., 2013). Combining same-sex and opposite-sex DZ twins in analyses has the potential to inflate heritability estimates if opposite-sex twins DZ twins are systematically less similar (because of sex effects) than same-sex DZ twins, as sex limited genetic effects will end up in a lower DZ correlation and thus an inflated genetic estimate. However, as many of the synthesized studies combined data from same-sex and opposite-sex DZ twins, I was not able to run sufficiently powered analyses using only estimates from same-sex twins or modelling same sex and opposite sex twins separately. I included the following types of estimate in the analyses: estimates from combined samples of same-sex and opposite-sex DZ twins (labelled DZ in Figures 4–13 and 20–35), DZ estimates from samples that did not specify whether twins were same-sex, opposite-sex or combined (labelled DZ in Figures 4A–13A and 20A–35A); estimates from same-sex DZ twins (labelled DZSS, DZF [if same-sex and female only], or DZM [if same-sex and male only] in Figures 4–13 and 20–35) and estimates from opposite-sex DZ twins (DZOS). For example, if a study only reported one estimate per phenotype from DZ twins I included only that one estimate, whereas if a study separately reported twin correlations for a phenotype from DZOS, DZM and DZF, I extracted and separately included all three of the estimates from that study. Dependency between estimates from the same study was taken into account in the multilevel models, which accounted for within- and between-cohort variance. The script for the meta-analyses is available on the Open

Science Framework:
https://osf.io/4y7z8/?view_only=8d98cb5ce4224e15a401fa5dd658e878.

A.1.5 Meta-Analytic SEM Models

After running the multilevel random effects meta-analyses, I calculated ACE estimates by running meta-analytic SEM models. These models were estimated using the correlations and squared standard errors (variances) from each of the two the multilevel meta-analyses (one to estimate MZ and one to estimate DZ), for each phenotype category, thus allowing parameters and their confidence intervals to be estimated taking account of between-study heterogeneity. Confidence intervals for the ACE parameter estimates were obtained using likelihood-based confidence interval estimation in OpenMx. The script for the meta-analytic twin modelling and confidence interval estimation is available on the Open Science Framework:
https://osf.io/4y7z8/?view_only=8d98cb5ce4224e15a401fa5dd658e878.

A1.1.6 Forest Plots

The script used to create the Forest plots displayed in (Figures 4A–13A, 20A–35A) is available on the Open Science Framework (https://osf.io/4y7z8/?view_only=8d98cb5ce4224e15a401fa5dd658e878). There are small discrepancies between the main results (presented in Table 1A and Tables 8A and 9A) and the pooled estimates in the forest plots because the main results were estimated based on all the pooled data whereas the forest plots were created using MZ and DZ subsets, separately. The main results, based on all the pooled data, provide the most precise estimates. The phenotype category with the greatest discrepancy between the two methods was ‘family relationships’.

A.2 Supplementary Results

A.2.1 Breakdown by Geographical Location

More than half (52.55%) the twin pairs were from European samples, around a quarter (24.09%) were from North American samples, 15.72% were based in Asia, 7.38% in Oceania and less than 1% were from African or South American samples (0.24% and 0.02%, respectively). For a breakdown of twin pairs by country, see Figure 2A, and by continent, see Figure 3A.

A.2.2 Exclusions from Meta-Analysis

Of the 33 papers excluded from the meta-analysis, 22 were excluded because they contained no novel data—i.e., they only contained relevant findings that were reported in one or more of the other included papers. Eighteen of the 377 phenotypes (4.8%) were excluded from the meta-analysis because they did not fit into a category of the ICF-CY (see final row of Table 6A), resulting in the exclusion of two papers. Fifty-nine phenotypes (15.6%) were excluded from the meta-analysis because they fell into a category of the ICF-CY that contained phenotypes gathered from fewer than five independent twin samples, resulting in the exclusion of nine papers. Extracted data, including the data excluded from the meta-analysis, is available on the Open Science Framework (https://osf.io/4y7z8/?view_only=8d98cb5ce4224e15a401fa5dd658e878).

A.2.3 Subcategory Analysis

The ten subcategories of the ICF-CY containing data from five or more independent samples were meta-analysed. Definitions for each subcategory can be found in the ICF-CY(World Health Organization, 2007). Lists of all phenotypes included in subcategories are displayed in Table 6A. Results from the meta-analyses of subcategories are reported in Table 8A. Forest plots for these analyses are reported in Figures 20A–29A.

Heritability. ‘Psychomotor control’ ($h^2_{pooled} = .61$; 95% CI [.26–.75] $p < .001$), and ‘regulating behaviors within interactions’ ($h^2_{pooled} = .58$; 95% CI [.18–.87] $p = .004$) had the highest heritability estimates. ‘Sustaining attention’, ‘organization of psychomotor functions’, ‘range of emotion’, ‘regulation of emotion’, and ‘social cues in relationships’ also had high heritability estimates (h^2_{pooled} range: .41–.50). Heritability estimates for ‘respect and warmth in relationships’, ‘acting in accordance with social rules’ and ‘expression of language’ were small-to-moderate and had confidence intervals that overlapped with 0 (h^2_{pooled} range: .20–.31).

Shared Environment: ‘Expression of language’ had the highest shared environmental estimate ($c^2_{pooled} = .59$; 95% CI [.34–.85] $p < .001$). ‘Respect and warmth in relationships’, ‘acting in accordance with social rules’, and ‘organization of psychomotor functions’ also had moderate-to-high shared environmental estimates (c^2_{pooled} range: .29–.42) with 95% confidence intervals that did not overlap with zero. Shared environmental estimates for ‘psychomotor control', 'sustaining attention', 'regulation of emotion', 'regulating behaviors', 'range of emotion', and 'social cues in relationships' were small and had 95% confidence intervals that overlapped with 0 (c^2_{pooled} range: .00–.18).

Nonshared Environment. Each of the ten phenotypic sub-categories had nonshared environmental estimates with 95% confidence intervals that did not overlap with zero (e^2_{pooled} range: .14–.47). The category with the highest nonshared environmental estimate was 'regulation of emotion' ($e^2_{pooled} = .47$; 95% CI [.35–.59] $p < .001$).

Heterogeneity. As displayed in Table 8A, sampling variance contributed a proportionally small amount to the total variance of each of the ten phenotypic sub-categories ($I^2_{Level\ 1}$ range: 0.21%–16.31%). Within-cohort heterogeneity contributed a substantial amount to the total variance in 'respect and warmth in relationships', 'range of emotion' and 'expression of language' ($I^2_{Level\ 2}$ range: 58.26%–81.71%) and between-cohort heterogeneity contributed a low or moderate amount to these outcomes ($I^2_{Level\ 3}$ range: 17.78%–41.53%). The remaining seven sub-categories each had substantial between-cohort heterogeneity ($I^2_{Level\ 3}$ = 51.30%–98.07%) and low or moderate within-cohort heterogeneity ($I^2_{Level\ 2}$ range: 0.00%–43.00%).

A.2.4 Analysis of Phenotypes by Category and Rater

Each estimate included in the meta-analysis was coded as: 'parent-rated', 'observer-rated', or 'other'. 'Other' included estimates where the rater was not specified, as well as objective measurement such as neuroimaging, measurement of birth weight/length, cortisol sampling and actigraphy. Twenty-seven independent cohorts contained estimates coded as 'other'. However, this was largely driven by the 'Growth maintenance functions' category, which contained data on anthropometric characteristics, the measurements for which were all coded as 'other'. Overall, there were more independent cohorts with parent-reported data ($k_{cohort} = 22$) than observer-rated data ($k_{cohort} = 12$). Examining the ICF-CY categories individually, in 7 out of the 10 categories there were more independent cohorts containing parent-report data than observer-rated data. It was only in the 'Basic cognitive functions' category that there were more cohorts reporting observer ratings than parent ratings. In 'Family relationships' there were an equal number of cohorts containing parent and observer reports. For a breakdown of the phenotypic categories by rater subgroup, see Table 7A. There is evidence that heritability estimates can differ by rater, thus I also conducted analyses of ICF-CY categories split by informant (parent or observer). Three of the ICF-CY categories contained data from five or more independent samples for both subgroups: (1) parent-reported and (2) observer-rated phenotypes. Data from each of these six subgroups (three parent-report and three observer-report) were meta-analysed. Results from these analyses are reported in Table 9A. Forest plots for the analyses are reported in Figures 30A–35A.

Heritability. The heritability point estimates for parent reports of ‘psychomotor functions’, ‘basic interpersonal interactions’ and ‘emotional functions’ were all high and had 95% confidence intervals that did not overlap with zero (h^2_{pooled} range: .49–.67). Conversely, heritability estimates for observer ratings of the same three phenotypic categories were smaller and had confidence intervals that overlapped with zero (h^2_{pooled} range: .20–.35).

Shared Environment. The shared environmental estimate for parent-reports of ‘emotional functions’ was small but had a confidence interval that overlapped with zero (c^2_{pooled} = .17; 95% CI [.01–.35] *p* = .033). The remaining two parent-report subgroups (‘psychomotor functions’ and ‘basic interpersonal interactions’) both had small shared environmental estimates with 95% confidence intervals overlapping with zero (c^2_{pooled}: .00 and .15, respectively) as did the three observer-reported subgroups c^2_{pooled} range: .12–.21). Comparable shared environment estimates in parent and observer ratings might be interpreted as indicating limited rater bias linked to shared environment. However, this comparison is likely confounded by other influences and is complicated by the fact that, in theory, studies may have had the same observer rating both twins (although, generally speaking, this is unusual, and it is hard to estimate in the present specific sample as this information was typically not provided by the included studies).

Nonshared Environment. Nonshared environmental estimates for parent and observer reports of ‘psychomotor functions’, ‘basic interpersonal interactions’ and ‘emotional functions’ were all statistically significant. The observer-report subgroups each had higher nonshared environmental estimates (e^2_{pooled} range: .45.47) than the parent-report subgroups (e^2_{pooled} range: .30.33).

Heterogeneity. As displayed in Table 9A, a proportionally small amount of the total variance was attributed to sampling variance in each of the six observer- or parent-report subgroups ($I^2_{\text{Level 1}}$ range = 0.88%–18.76%). Observer ratings of ‘psychomotor functions’ and ‘emotional functions’, and parent ratings of ‘emotional functions’, each had moderate to substantial within-cohort heterogeneity ($I^2_{\text{Level 2}}$ range = 43.14%–94.31%). Conversely, parent reports of ‘psychomotor functions’ and ‘basic interpersonal interactions’, and observer reports of ‘basic interpersonal interactions’, each had low within-cohort heterogeneity ($I^2_{\text{Level 2}}$ range = 2.92%–32.18%). Parent reported ‘emotional functions’ had between-cohort heterogeneity of 0% and the remaining five observer- and parent-reported categories all had moderate to substantial between cohort heterogeneity ($I^2_{\text{Level 3}}$ = 40.38%–96.20%).

A.2.5 Publication Bias

Publication bias is present when the likelihood of a finding being published is influenced by the finding itself. For example, if statistically significant findings are more likely to be published than findings that are not statistically significant. O examined publication bias via two methods. First, I created funnel plots, plotting effect sizes against standard errors. These plots are reported in Figures 14A–18A. As larger studies typically have lower standard errors than smaller studies, larger studies should appear towards the top of the plot and smaller ones towards the bottom. It is assumed that, in the absence of publication bias, smaller studies with higher standard errors will be scattered symmetrically across the bottom of the plot. In the presence of publication bias it is assumed that plots will be skewed, displaying asymmetry, e.g., there will be more studies with small sample sizes reporting positive than negative results. Based on visual inspection of the funnel plots of rMZ and rDZ in Figures 14A and 15A, there appears to be some asymmetry. Specifically, larger studies tended to publish findings with stronger correlations and smaller studies tended to publish findings with weaker correlations. As displayed in Figure 15A, estimates of h^2 largely displayed no patterns of asymmetry, although for ‘psychomotor functions’ and ‘growth maintenance functions’ the smaller studies tended to report lower h^2 estimates. As shown in Figures 17A and 18A, for most of the phenotypic domains, estimates of c^2 and e^2 echoed the same patterns of asymmetry seen in correlations—with smaller studies tending to publish smaller effects. As the interpretation of funnel plots depends on visual examination and is thus somewhat subjective, I also ran the Egger's regression test of publication bias. The Egger's test regresses effect sizes on their standard errors, weighted by their inverse variances. In the absence of publication bias, the regression intercept is expected to be zero and in the presence of publication bias, it is expected to be significantly different from zero ($p < .05$). Results of Egger's regression test are presented in Tables 10A and 11A. Echoing the patterns of skewness in the funnel plots, there was evidence of publication bias for all correlations by category ($p < .001$), apart from *r*DZ in 'attention functions' and 'family relationships' ($p = .061$ and .062). Results suggested potential publication bias in heritability estimates for 'psychomotor' and 'emotional' functions, 'basic interpersonal interactions' and 'complex interpersonal interactions' (*p* range: .013 to < .001) but not 'sleep', 'attention' or 'growth maintenance' functions (*p* range: .072–.385). Findings indicated publication bias for all shared environment estimates ('emotional functions', $p = 0.08$ and all others $p < .001$), apart from 'attention functions' ($p = .402$). Results suggested publication bias in nonshared environment estimates of 'attention functions' and 'basic interpersonal interactions' ($p = .036$, .027, respectively) but not 'sleep', 'psychomotor', 'emotional' or 'growth maintenance' functions or 'complex interpersonal interactions' (*p* range: .099–.494). The possible publication bias observed in the funnel plots and Egger's test results was in the opposite direction to the publication bias that is typically hypothesized in research. Usually, it is expected that publication bias will result in more smaller studies publishing large effect sizes due to a bias to publish positive or favourable findings. However, in this case, smaller studies tended to publish findings with smaller effect sizes. These unusual results may reflect that fact that statically significant within-twin correlations and estimates of heritability, shared and non-shared environmental influence do not represent "favourable" results in the same way that statistically significant results might in, for example, intervention research. For example, it is not the statistical significance of a within-twin correlation but, rather, the relationship between within-twin MZ correlations and within-twin DZ correlations that is drawn upon to estimate the sizes of the three variance components. The three variance components each represent the proportion of variation linked to individual differences in a trait.

Individually, none of them can be below 0 or above 100 and collectively they must sum to 100. Consequently, these estimates represent something quite different to effect sizes in other research contexts. Additionally, while funnel plot asymmetry and significant Egger's test results indicate possible publication bias, they could be driven by alternative explanations. One particularly plausible possibility, given the highly heterogeneous findings, is that the asymmetry was influenced by between-study heterogeneity, which can also induce skewness in funnel plots.(Page, Sterne, et al., 2021)

Table 1A. Database and Reference List Searches

Date	Database	Search terms	N papers	N papers (no duplicates)	Includes (T&A)	Includes (FT)
30/11/18	PubMed	(twin*[Title/Abstract] AND (gene[Title/Abstract] OR genome[Title/Abstract] OR genetic* [Title/Abstract] OR herita* [Title/Abstract] OR environment* [Title/Abstract]) AND (infan*[Title/Abstract] OR early [Title/Abstract]))	2735	2727	199	105
30/11/18	PsycINFO	(twin* and (gene or genome or genetic* or herita* or environment*) and (infan* or early)).ab.	1291	504	47	13
05/02/20	PubMed (from 1/10/18 onwards)	(twin*[Title/Abstract] AND (gene[Title/Abstract] OR genome[Title/Abstract] OR genetic* [Title/Abstract] OR herita* [Title/Abstract] OR environment*[Title/Abstract]) AND (infan*[Title/Abstract] OR early[Title/Abstract]))	315	230	5	1
05/02/20	PsycINFO (From 2018 onwards)	(twin* and (gene or genome or genetic* or herita* or environment*) and (infan* or early)).ab.	120	28	0	0
11/02/21	PubMed (from 1/2/20 onwards)	(twin*[Title/Abstract] AND (gene[Title/Abstract] OR genome[Title/Abstract] OR genetic* [Title/Abstract] OR herita* [Title/Abstract] OR environment*[Title/Abstract]) AND (infan*[Title/Abstract] OR early[Title/Abstract]))	190	179	4	2
11/02/21	PsycINFO (From 2020 onwards)	(twin* and (gene or genome or genetic* or herita* or environment*) and (infan* or early)).ab.	24	7	2	0
n/a	Reference list searching	n/a	372	98	98	18
			5047	3773	355	139

Note. N = Number of. T&A = After screening by title and abstract. FT = After screening by full text.

Table 2A. Inclusion and Exclusion Criteria

	Inclusion criteria	Exclusion criteria
Publication type	➢ Peer-reviewed journals ➢ In English	➢ Not subject to peer review ➢ Not published in English
Study type	➢ Twin studies	➢ Not reporting twin data ➢ Syntheses of previously published twin data (e.g., reviews/meta-analyses)
Population	➢ MZ and DZ twins ➢ Infants with mean age ≤ 2	➢ Mean age > 2 years ➢ MZ twins only ➢ Case studies
Outcome	➢ Psychological and developmental traits and milestones ➢ MZ/DZ correlations ➢ ACE estimates based only on data from MZ/DZ twins	➢ Not on psychological and developmental traits and milestones ➢ No heritability estimates, twin correlations, or concordances available ➢ Twin data not possible to separate from non-twin data ➢ Multivariate or gene-environment interaction analyses where univariate estimates not possible to extract

Note. MZ = monozygotic. DZ = dizygotic

Table 3A. Information Extracted from Included Papers

Extracted data	Description
Study	Overarching twin study/sample/cohort
Country	Country in which twin population was based
Sex	Whether the sample included male and/or female twins
Phenotype	The examined trait as labelled in the paper
Age	In months, at which data was collected
Rater	Who rated the reported data (e.g., parent, observer)
Measure	How the phenotype was measured
MZ/DZ correlations	Within-twin monozygotic twin correlations and dizygotic twin correlations (based on samples of same-sex DZ twins, opposite-sex DZ twins and combined samples)
ACE estimates	Estimates of heritability (A, or h^2) and shared (C, or c^2) and nonshared (E, or e^2) environmental influence
Number of twin pairs	Entered for each extracted correlation/variance component
Method for estimating variance components	Method used to calculate ACE estimates
Continuous or dichotomous data	Was the measure used to capture the trait continuous or dichotomous
Concordant/discordant pairs	If a dichotomous measure, rates of concordance/discordance for the trait
Prevalence	If a dichotomous measure, trait prevalence rates in the twin sample being examined

Table 4A. Papers Identified in the Systematic Literature Search, Presented Alphabetically by First Author

Paper	Status	Study	Country	Category (using codes from the ICF-CY[a])	Phenotypes Extracted from Paper
Akerman & Fischbein (1992)	Included in meta-analysis	Project Metropolitan	Sweden	b560 Growth maintenance functions[b]	Height Weight
Ando et al. (2006)	Included in meta-analysis	Tokyo Twin Cohort Project	Japan	b134 Sleep functions b560 Growth maintenance functions d550 Eating[c] d710 Basic interpersonal interactions	Chest circumference Head circumference Height Milk consumption Mimic, point gazing, joint attention Rhythmicity Time to fall asleep Weight
Bakermans-Kranenburg et al. (2004)	Included in meta-analysis	Netherlands Twin Register	The Netherlands	d760 Family relationships	Dependency (to father) Infant-father attachment security
Beaver et al. (2014)	Included in meta-analysis	Early Childhood Longitudinal Study-Birth Cohort	United States	b167 Mental functions of language	Expressive vocabulary
Bishop et al. (2003b)	Excluded from meta-analysis (duplicate data)	The Colorado Twin Registry: Longitudinal Twin Sample	United States	b163 Basic cognitive functions	General cognitive ability

---

[a] ICF-CY = International Classification of Functioning, Disability and Health, Children and Youth Version.[4] Definitions for each of the categories and subcategories can be found in the cited ICF-CY manual: https://apps.who.int/iris/handle/10665/43737.

[b] The prefix *b* is given to coded items within the ICF-CY component of Body Functions.

[c] The prefix *d* is given to items in the component of Activities and Participation.

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Bokhorst et al. (2003)	Included in meta-analysis	Netherlands Twin Register and Multiple Births Foundation of Queen Charlotte's and Chelsea Hospital in London	The Netherlands; UK	b152 Emotional functions d760 Family relationships	Attachment disorganization Attachment security Temperamental reactivity
Boomsma et al. (1992)	Excluded from meta-analysis (duplicate data)	Netherlands Twin Register	The Netherlands	b560 Growth maintenance functions	Height Weight
Brant et al. (2009)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b163 Basic cognitive functions	General cognitive ability
Brescianini et al. (2011)	Included in meta-analysis	Italian Twin Register	Italy	b134 Sleep functions	Cosleeping Diurnal sleep duration Night awakenings Nocturnal sleep duration
Brescianini et al. (2012)	Included in meta-analysis	Italian Twin Register	Italy	b560 Growth maintenance functions	Weight gain
Caramaschi et al. (2012)	Excluded from meta-analysis (phenotype[s] could not be categorized)	Quebec Newborn Twin Study	Canada	N/A	Testosterone

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Chen et al. (1990a)	Included in meta-analysis	Taipei City Teaching Hospitals Twin Study	Taiwan	b125 Dispositions and intra-personal functions b140 Attention functions b147 Psychomotor functions b152 Emotional functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Activity level Adaptability Approach/withdrawal Attention/persistence Distractibility Intensity of reaction Quality of mood Rhythmicity Threshold of responsiveness
Chen et al. (1990b)	Included in meta-analysis	Taipei City Teaching Hospitals Twin Study	Taiwan	b560 Growth maintenance functions	Arm circumference Chest circumference Head circumference Height Weight
Cherny et al. (1992)	Excluded from meta-analysis (duplicate data)	The Colorado Twin Registry: Longitudinal Twin Sample	United States	b163 Basic cognitive functions	Cognitive ability
Cherny et al. (1994a)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	d710 Basic interpersonal interactions	Shyness
Cherny et al. (1994b)	Excluded from meta-analysis (duplicate data)	The Colorado Twin Registry: Longitudinal Twin Sample	United States	b163 Basic cognitive functions	Cognitive ability
Custodio et al. (2007)	Included in meta-analysis	University of Sao Paulo - Longitudinal Twin Study	Brazil	b134 Sleep functions	Emergence of the cortisol circadian rhythm

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Dale et al. (2000)	Included in meta-analysis	Twins Early Development Study	UK	b167 Mental functions of language	Grammar Vocabulary
	Included in meta-analysis	Louisville Twin Study	United States	b140 Attention functions b147 Psychomotor functions d710 Basic interpersonal interactions	Activity level Affect-extraversion Task orientation
	Included in meta-analysis	Colorado Twin Registry: Twin Infant Project	United States	b140 Attention functions b152 Emotional functions b310 Voice functions d710 Basic interpersonal interactions	Affection for mother Enthusiasm for interaction with mother Negative affect Task orientation Vocalize Watch mother
	Excluded from meta-analysis (duplicate data)	The Colorado Twin Registry: Longitudinal Twin Sample	United States	d710 Basic interpersonal interactions	Behavioral Inhibition
	Excluded from meta-analysis (duplicate data)	Twins Early Development Study	UK	b167 Mental functions of language	Grammar Vocabulary
	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b167 Mental functions of language d720 Complex interpersonal interactions	Expressive vocabulary Physical aggression
	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b134 Sleep functions b167 Mental functions of language	Ratio of day/night sleep duration Vocabulary
	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b560 Growth maintenance functions	Height Weight

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Dubois et al. (2012)	Included in meta-analysis	Quebec Newborn Twin Study, Child and Adolescent Twin Study in Sweden, Twin Study of Child and Adolescent Development, Danish Twin Registry, and Brisbane Longitudinal Twin Study	Canada, Sweden, Denmark, Australia	b560 Growth maintenance functions	BMI Height Weight
Emde et al. (1992)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b140 Attention functions b144 Memory functions b147 Psychomotor functions b152 Emotional functions b164 Higher-level cognitive functions b167 Mental functions of language d710 Basic interpersonal interactions	Activity Affect Attention/persistence Behavioral Inhibition Categorization Emotionality Empathy Expressive language Frustration Memory for location Negative affect Negative hedonic tone Overall mood Positive affect Positive hedonic tone Receptive language Shyness Sociability Task orientation Word comprehension

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Finkel et al. (2000)	Included in meta-analysis	Louisville Twin Study	United States	d760 Family relationships	Attachment
Fisher et al. (2012)	Included in meta-analysis	Gemini Study	UK	b134 Sleep functions	Daytime nap duration Night awakenings Nighttime sleep duration Wake time
Flom & Saudino (2017)	Excluded from meta-analysis (duplicate data)	Boston University Twin Project	United States	d710 Basic interpersonal interactions	Callous unemotional behavior
Flom & Saudino (2018)	Included in meta-analysis	Boston University Twin Project	United States	b140 Attention functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions	ADHD Callous unemotional behavior ODD
Flom et al. (2019)	Included in meta-analysis	Boston University Twin Project	United States	d710 Basic interpersonal interactions	Callous unemotional traits
Forget-Dubois et al. (2007)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	d720 Complex interpersonal interactions	Difficult temperament Disruptive behavior
Friedman et al. (2011)	Excluded from meta-analysis (phenotype in category containing < 5 independent samples)	Colorado Twin Registry: Longitudinal Twin Sample	United States	b164 Higher-level cognitive functions	Self-restraint

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Fujisawa et al. (2012)	Included in meta-analysis	Tokyo Twin Cohort Project	Japan	b560 Growth maintenance functions d710 Basic interpersonal interactions	Head circumference Head circumference growth Sociocognitive abilities
Gagne & Goldsmith (2011)	Included in meta-analysis	Wisconsin Twin Panel	United States	b152 Emotional functions	Anger Distress to limitations
Gagne & Saudino (2010)	Included in meta-analysis	Boston University Twin Project	United States	b164 Higher-level cognitive functions	Inhibitory control
Gagne et al. (2011)	Included in meta-analysis	Boston University Twin Project	United States	b140 Attention functions b164 Higher-level cognitive functions d720 Complex interpersonal interactions	ADHD Externalizing Inhibitory control
Gagne & Saudino (2016)	Excluded from meta-analysis (duplicate data)	Boston University Twin Project	United States	b164 Higher-level cognitive functions	Inhibitory control
Gagne et al. (2020)	Excluded from meta-analysis (duplicate data and phenotype in category containing < 5 independent samples)	Boston University Twin Project	United States	b140 Attention functions b164 Higher-level cognitive functions	ADHD Inhibitory control
Galsworthy et al. (2000)	Excluded from meta-analysis (duplicate data)	Twins Early Development Study	UK	b163 Basic cognitive functions b167 Mental functions of language	Non-verbal cognitive development Verbal ability

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
German et al. (2015)	Included in meta-analysis	Childcare Centers of Tel Aviv and Haifa	Israel	b560 Growth maintenance functions	Age at transition to childhood
Gilmore et al. (2010)	Excluded from meta-analysis (phenotypes in category containing < 5 independent samples)	UNC Early Brain Development Study	United States	s110 Structure of brain[d]	Cerebellum Corpus Callosum Cortical grey matter Cortical unmyelinated white matter Frontal grey matter Frontal unmyelinated white matter Intracranial volume Lateral Ventricles Left hemisphere grey matter Left hemisphere total Left hemisphere unmyelinated white matter Occipital grey matter Occipital unmyelinated white matter Parietal grey matter Parietal unmyelinated white matter Prefrontal grey matter Prefrontal unmyelinated white matter Right hemisphere grey matter Right hemisphere total Right hemisphere unmyelinated white matter Subcortical grey matter Total cerebrospinal fluid Total early myelinated white matter Total Frontal Total grey matter Total Occipital Total Parietal Total Prefrontal Total unmyelinated white matter

---

[d] The prefix *s* is given to items in the ICF-CY component of Body Structures.

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Goetghebuer et al. (2003)	Included in meta-analysis	Twin study in The Gambia	The Gambia	b147 Psychomotor functions b560 Growth maintenance functions	Crawl Length Maintain head Roll over Sitting Sitting without support Stand holding on something Take two steps Walk holding on something Weight
Goldsmith & Gottesman (1981)	Included in meta-analysis	Collaborative Perinatal Project	United States	b125 Dispositions and intra-personal functions b140 Attention functions b147 Psychomotor functions b560 Growth maintenance functions d710 Basic interpersonal interactions	Active manipulation Activity level Degree of social acceptance of examiner Degree of social contact with mother Interest in persons Interest in/responsiveness to people Physical development Pursuit persistence Response duration Speed of response Vigorous activity vs. psychomotor passivity
Goldsmith et al. (1999)	Included in meta-analysis	Pooled Sample of twins from Oregon, Washington, Colorado, Texas, and Wisconsin	United States	b140 Attention functions b147 Psychomotor functions b152 Emotional functions	Activity level Distress to limitations Distress to novelty Duration of orienting Negative affect Positive affect Resistance to soothing Smiling and laughter

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Hawks & Marrus (2018)	Included in meta-analysis	Early Reciprocal Social Behavior Study	United States	d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Behavior problems Competence Reciprocal social behavior
Herle et al. (2018)	Excluded from meta-analysis (phenotype in category containing < 5 independent samples)	Gemini Study	UK	d550 Eating	Emotional overeating
Hur (2005)	Excluded from meta-analysis (duplicate data)	The South Korean Twin Registry (Previously the Seoul Twin Family Study)	South Korea	b560 Growth maintenance functions	Weight

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Hur et al. (2005)	Included in meta-analysis	Brisbane Adolescent Twin Study Japanese pooled sample: (1) Registry of twins recruited in Japan from associations for parents of multiples; (2) Registry of twins who applied for the secondary school attached to the Faculty of Education at the University of Tokyo between 1981 and 2003. Minnesota Twin Family Study South Korean pooled sample: (1) Twins born in two South Korean hospitals (1998–2003); (2) Seoul Twin Family Study Netherlands Twin Register	Australia, Japan, United States, South Korea, The Netherlands	b560 Growth maintenance functions	Weight

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Ilott et al. (2010a)	Included in meta-analysis	Boston University Twin Project	United States	b140 Attention functions b147 Psychomotor functions	Activity Level ADHD
Ilott et al. (2010b)	Excluded from meta-analysis (duplicate data)	Boston University Twin Project	United States	b140 Attention functions	ADHD
Jackson (2016)	Included in meta-analysis	Early Childhood Longitudinal Study-Birth Cohort	United States	b152 Emotional functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Attachment security Avoids others/not sociable Comfortable cuddly Cooperative Demanding/angry Dependency Enjoys company Independent Moody/unusual Seeks attention Upset by separation
Jha et al. (2018)	Excluded from meta-analysis (phenotypes in category containing < 5 independent samples)	UNC Early Brain Development Study	United States	s110 Structure of brain	Cortical surface area Cortical thickness Intracranial volume
Johnson et al. (2011)	Included in meta-analysis	Gemini Study	UK	b560 Growth maintenance functions	Size (weight) Tempo (weight) Weight Weight change Weight velocity

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Knafo & Plomin (2006)	Included in meta-analysis	Twins Early Development Study	UK	d710 Basic interpersonal interactions	Prosocial behavior
Koeppen-Schomerus et al. (2003)	Included in meta-analysis	Twins Early Development Study	UK	b163 Basic cognitive functions b167 Mental functions of language d720 Complex interpersonal interactions	Behavior problems General cognitive ability Nonverbal cognitive ability Verbal ability
Kuntsi et al. (2005)	Included in meta-analysis	Twins Early Development Study	UK	b147 Psychomotor functions	Hyperactivity
Lacourse et al. 2014	Included in meta-analysis	Quebec Newborn Twin Study	Canada	d720 Complex interpersonal interactions	Physical aggression
Levine et al. (1987)	Included in meta-analysis	Jackson Memorial Hospital/University of Miami Medical Center	United States	b560 Growth maintenance functions	Length Weight
Liu et al. (2015)	Included in meta-analysis	Chinese National Twin Registry	China	b560 Growth maintenance functions	BMI Height Weight

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Livshits et al. (2000)	Included in meta-analysis	Childcare Centers of Tel Aviv and Haifa	Israel	b560 Growth maintenance functions	Head circumference Head circumference (growth curve A) Head circumference (growth curve B) Head circumference (growth curve C) Height Height (growth curve A) Height (growth curve B) Height (growth curve C) Weight Weight (growth curve A) Weight (growth curve B) Weight (growth curve C)
Llewellyn et al. (2010)	Excluded from meta-analysis (phenotypes in category containing < 5 independent samples)	Gemini Study	UK	d550 Eating	Enjoyment of food Food responsiveness Satiety responsiveness Slowness in eating
Marrus et al. (2015)	Included in meta-analysis	Early Reciprocal Social Behavior Study	United States	d710 Basic interpersonal interactions	Reciprocal social behavior Restrictive repetitive behavior
Marrus et al. (2018)	Included in meta-analysis	Early Reciprocal Social Behavior Study	United States	d710 Basic interpersonal interactions	Reciprocal social behavior
Marrus et al. (2020)	Included in meta-analysis	Early Reciprocal Social Behavior Study	United States	d710 Basic interpersonal interactions	Functional communication Restrictive repetitive behavior Social avoidance Social motivation Social orienting

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Matheny (1980)	Included in meta-analysis	Louisville Twin Study	United States	b140 Attention functions b147 Psychomotor functions b279 Additional sensory functions, other specified and unspecified d710 Basic interpersonal interactions	Activity level Activity Level Affect-extraversion Auditory-visual Motor development Task orientation
Matheny (1983)	Included in meta-analysis	Louisville Twin Study	United States	b140 Attention functions b147 Psychomotor functions	Activity level Task orientation
Matheny (1984)	Excluded from meta-analysis (phenotype[s] could not be categorized)	Louisville Twin Study	United States	N/A	General temperament
Matheny (1989)	Included in meta-analysis	Louisville Twin Study	United States	b152 Emotional functions d710 Basic interpersonal interactions	Approach/withdrawal Behavioral inhibition Emotional tone Fearfulness

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Matheny et al. (1976)	Included in meta-analysis	Louisville Twin Study	United States	b140 Attention functions b147 Psychomotor functions b152 Emotional functions b163 Basic cognitive functions b310 Voice functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Attention Banging Cooperative Emotional tone Endurance Energy Extraversion Fearfulness Fine motor Goal directedness Gross motor Listening Looking Manipulating Mouthing: pacifier Mouthing: thumb Mouthing: toys Object orientation Primary cognition Reactivity Tension Vocalize
Micalizzi et al. (2016)	Included in meta-analysis	Boston University Twin Project	United States	b152 Emotional functions	Affective problems Autistic-like traits
Micalizzi et al. (2017)	Included in meta-analysis	Boston University Twin Project	United States	d720 Complex interpersonal interactions	Difficult temperament
Mook-Kanamori et al. (2012)	Included in meta-analysis	Netherlands Twin Register	The Netherlands	b560 Growth maintenance functions	Height Length Weight
Nguyen et al. (2008)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b134 Sleep functions	Sleep terrors
Nichols et al. (1974)	Included in meta-analysis	Collaborative Perinatal Project	United States	b163 Basic cognitive functions	Mental development

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Orekhova et al. (2003)	Included in meta-analysis	Moscow City Twin Sample	Russia	b140 Attention functions	Alpha frequency during darkness EEG mu gravity frequency during visual attention EEG spectral amplitude during visual attention: Alpha AF3 EEG spectral amplitude during visual attention: Alpha AF4 EEG spectral amplitude during visual attention: Alpha F7 EEG spectral amplitude during visual attention: Alpha F8 EEG spectral amplitude during visual attention: Alpha FC3 EEG spectral amplitude during visual attention: Alpha FC4 EEG spectral amplitude during visual attention: Alpha O1 EEG spectral amplitude during visual attention: Alpha O2 EEG spectral amplitude during visual attention: Alpha PO3 EEG spectral amplitude during visual attention: Alpha PO4 EEG spectral amplitude during visual attention: Alpha T5 EEG spectral amplitude during visual attention: Alpha T6 EEG spectral amplitude during visual attention: Delta AF3 EEG spectral amplitude during visual attention: Delta AF4 EEG spectral amplitude during visual attention: Delta F7 EEG spectral amplitude during visual attention: Delta F8 EEG spectral amplitude during visual attention: Delta FC3 EEG spectral amplitude during visual attention: Delta FC4 EEG spectral amplitude during visual attention: Delta O1 EEG spectral amplitude during visual attention: Delta O2 EEG spectral amplitude during visual attention: Delta PO3 EEG spectral amplitude during visual attention: Delta PO4 EEG spectral amplitude during visual attention: Delta T5 EEG spectral amplitude during visual attention: Delta T6 EEG spectral amplitude during visual attention: Theta AF3 EEG spectral amplitude during visual attention: Theta AF4 EEG spectral amplitude during visual attention: Theta F7 EEG spectral amplitude during visual attention: Theta F8 EEG spectral amplitude during visual attention: Theta FC3 EEG spectral amplitude during visual attention: Theta FC4 EEG spectral amplitude during visual attention: Theta O1 EEG spectral amplitude during visual attention: Theta O2 EEG spectral amplitude during visual attention: Theta PO3 EEG spectral amplitude during visual attention: Theta PO4 EEG spectral amplitude during visual attention: Theta T5 EEG spectral amplitude during visual attention: Theta T6 Spectral amplitude during darkness

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Ouellet-Morin et al. (2008)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b152 Emotional functions	Cortisol reactivity
Ouellet-Morin et al. (2009)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b134 Sleep functions	Cortisol secretion in morning Cortisol secretion on awakening
Peter et al. (1999)	Included in meta-analysis	Childcare Centers of Tel Aviv and Haifa	Israel	b147 Psychomotor functions	Pulling up to a standing position Sitting up Turning over Walking five steps
Petitclerc et al. (2011)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	d720 Complex interpersonal interactions	Disregard for rules
Pimpin et al. (2013)	Excluded from meta-analysis (phenotypes in category containing < 5 independent samples)	Gemini Study	UK	d550 Eating	Carbohydrate intake Energy intake Fat intake Food weight Protein intake
Planalp et al. (2017)	Included in meta-analysis	Waisman Center Birth to 3 year project	United States	b152 Emotional functions	Positive affect Smiling and laughter

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Plomin & Rowe (1979)	Included in meta-analysis	Denver Twin Sample	United States	d710 Basic interpersonal interactions d760 Family relationships	Approaching mother Approaching stranger Cuddliness with mother Cuddliness with stranger Difference of response between mother and stranger: approach Difference of response between mother and stranger: cuddliness Difference of response between mother and stranger: looking Difference of response between mother and stranger: positive vocalizations Difference of response between mother and stranger: proximity Difference of response between mother and stranger: quality of play Difference of response between mother and stranger: smiling Difference of response between mother and stranger: touches Latency to approach stranger Looking at mother Looking at stranger Positive vocalization to mother Positive vocalization to stranger Proximity to mother Proximity to stranger Quality of play with mother Quality of play with stranger Separation distress Smiling at mother Smiling at stranger Touching mother Touching stranger

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Plomin et al. (1993)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b140 Attention functions b144 Memory functions b147 Psychomotor functions b152 Emotional functions b163 Basic cognitive functions b164 Higher-level cognitive functions b167 Mental functions of language d710 Basic interpersonal interactions	Activity level Affect Attention/persistence Behavioral Inhibition Behavioral inhibition Categorization Emotionality Empathy Expressive language General cognitive ability Memory for Location Negative affect Overall mood Positive affect Positive hedonic tone Reactivity Receptive language Shyness Sociability Task orientation Word comprehension
Price et al. (2000)	Excluded from meta-analysis (duplicate data)	Twins Early Development Study	UK	b163 Basic cognitive functions b167 Mental functions of language	Nonverbal cognitive development Verbal ability
Price et al. (2005)	Included in meta-analysis	Twins Early Development Study	UK	b140 Attention functions	ADHD

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Pushina et al. (2005)	Excluded from meta-analysis (phenotype in category containing < 5 independent samples)	Moscow City twin sample	Russia	b144 Memory functions	Working memory
Reznick et al (1997)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b140 Attention functions b144 Memory functions b163 Basic cognitive functions b167 Mental functions of language	Expressive language Memory for Location Nonverbal Receptive language Verbal expressive Verbal receptive Visual attentiveness Word comprehension
Rhee et al. (2007)	Excluded from meta-analysis (duplicate data)	The Colorado Twin Registry: Longitudinal Twin Sample	United States	b152 Emotional functions	Emotionality
Rhee et al. (2012)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b152 Emotional functions	Negative affect
Rhee et al. (2013)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	d710 Basic interpersonal interactions	Observed Disregard
Rhee et al. (2016)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	d710 Basic interpersonal interactions	Disregard for others

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Rice et al. (2014)	Included in meta-analysis	Western Australian Twin Registry	Australia	b167 Mental functions of language	Combining words Late language acquisition Use of finiteness grammatical markers Words Produced
Riese (1990a)	Included in meta-analysis	Louisville Twin Study	United States	b147 Psychomotor functions b152 Emotional functions b560 Growth maintenance functions d710 Basic interpersonal interactions	Activity-awake Activity-sleep Irritability Reactivity Reinforcement Value Resistance to soothing Weight
Riese (1990b)	Included in meta-analysis	Louisville Twin Study	United States	b147 Psychomotor functions b152 Emotional functions	Activity-awake Activity-sleep Irritability Reactivity Resistance to soothing
Robinson et al. (1992)	Excluded from meta-analysis (duplicate data)	The Colorado Twin Registry: Longitudinal Twin Sample	United States	d710 Basic interpersonal interactions	Behavioral Inhibition
Roisman & Fraley (2006)	Included in meta-analysis	Early Childhood Longitudinal Study-Birth Cohort	United States	b152 Emotional functions	Fussiness and demanding behavior Positive and negative affect
Roisman & Fraley (2008)	Excluded from meta-analysis (duplicate data)	Early Childhood Longitudinal Study-Birth Cohort	United States	d760 Family relationships	Attachment security Temperamental dependency

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Ronald et al. (2010)	Included in meta-analysis	Boston University Twin Project	United States	b140 Attention functions d710 Basic interpersonal interactions	ADHD Autistic-like traits Non-social autistic-like traits Social autistic-like traits
Saudino (2012)	Included in meta-analysis	Boston University Twin Project	United States	b147 Psychomotor functions	Activity level
Saudino & Zapfe (2008)	Excluded from meta-analysis (duplicate data)	Boston University Twin Project	United States	b147 Psychomotor functions	Activity Level Home Activity Level Lab Activity Level Play
Saudino & Eaton (1991)	Included in meta-analysis	Manitoba Twin Study	Canada	b147 Psychomotor functions b560 Growth maintenance functions	Activity level Head circumference Length Motor development Ponderal index Weight
Saudino et al. (1996)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b140 Attention functions b147 Psychomotor functions d710 Basic interpersonal interactions	Activity level Affect-extraversion Task orientation

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Saudino et al. (2008)	Included in meta-analysis	Jumeaux et plus	France	b134 Sleep functions b140 Attention functions b147 Psychomotor functions b152 Emotional functions b279 Additional sensory functions, other specified and unspecified d550 Eating d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Activity/Impulsivity Aggression/Defiance Attention Atypical Index Competence Compliance Depression withdrawal Dysregulation Eating problems Empathy Externalizing General anxiety Imitation/Play Inhibition to novelty Internalizing Maladaptive behaviors Mastery Motivation Negative affect Peer aggression Prosocial peer relations Sensory Sensitivity Separation distress Sleep problems Social relatedness
Saudino et al. (2018)	Included in meta-analysis	Boston University Twin Project	United States	b140 Attention functions b147 Psychomotor functions	Activity level Attention problems
Schmitz et al. (1999)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	b152 Emotional functions d710 Basic interpersonal interactions	Emotionality Shyness
Schumann et al. (2017)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b152 Emotional functions	Negative affect

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Silberg et al. (2005)	Included in meta-analysis	Puerto Rican Infant Twin Study	United States	d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Difficult temperament Unadaptability Unsociability
Silberg et al. (2015)	Included in meta-analysis	Puerto Rican Infant Twin Study	United States	d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Difficultness Inhibition Resistance to control Sociability
Silventoinen et al. (2007)	Included in meta-analysis	Swedish Young Male Twins Study	Sweden	b560 Growth maintenance functions	BMI
Silventoinen et al. (2008)	Included in meta-analysis	Swedish Young Male Twins Study	Sweden	b560 Growth maintenance functions	Height
Silventoinen et al. (2011a)	Included in meta-analysis	West Japan Twins and Higher Order Multiple Births Registry	Japan	b560 Growth maintenance functions	Head circumference
Silventoinen et al. (2011b)	Included in meta-analysis	West Japan Twins and Higher Order Multiple Births Registry	Japan	b560 Growth maintenance functions	Height
Silventoinen et al. (2012)	Included in meta-analysis	West Japan Twins and Higher Order Multiple Births Registry	Japan	b560 Growth maintenance functions	Chest Circumference Chest circumference increase

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Silventoinen et al. (2016)	Included in meta-analysis	The CODATwins Project: Boston University Twin Project; Gemini Study; Guinea-Bissau Twin Study; Hungarian Twin Registry; Italian Twin Registry; Japanese Twin Cohort; Michigan State University Twin Registry; Mongolian Twin Registry; Netherlands Twin Registry; Peri/Postnatal Epigenetic Twins Study; Quebec Newborn Twin Study; Swedish Young Male Twins Study; Twins Early Development Study; West Japan Twins and Higher Order Multiple Births Registry	Canada, Guinea-Bissau, Hungary, Italy, Japan, Mongolia, Sweden, The Netherlands, United Kingdom, United States	b560 Growth maintenance functions	BMI

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Smit et al. (2010)	Included in meta-analysis	Netherlands Twin Register	The Netherlands	b560 Growth maintenance functions	Head circumference
Smith et al. (2012)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample	United States	d710 Basic interpersonal interactions	Behavioral inhibition
Smith et al. (2017a)	Excluded from meta-analysis (phenotypes in category containing < 5 independent samples)	Gemini Study	UK	d550 Eating	Food fussiness Food neophobia
Smith et al. (2017b)	Included in meta-analysis	Gemini Study	UK	b147 Psychomotor functions	Activity level First Crawl First Sit First Steps
Soussignan et al. (2009)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	d710 Basic interpersonal interactions	Emotional response to social stimuli Gaze aversion Motor activity during social stimuli Self-contact during social stimuli Social gaze
Spinath et al. (2003)	Included in meta-analysis	Twins Early Development Study	UK	b163 Basic cognitive functions	General cognitive ability
Stevenson & Fielding (1985)	Included in meta-analysis	Department of Psychology, University of Surrey	UK	b147 Psychomotor functions b152 Emotional functions b164 Higher-level cognitive functions d710 Basic interpersonal interactions	Activity level Emotionality Impulsivity Sociability

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Stroganova et al. (2000)	Included in meta-analysis	Moscow City Twin Sample	Russia	b140 Attention functions b152 Emotional functions b164 Higher-level cognitive functions b279 Additional sensory functions, other specified and unspecified d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Aggression towards father Aggression towards mother Autonomy Control Defensive reactions Dependence on mother Fear High tension Imitation of father Imitation of mother Love for father Love for mother Low tension Moderate tension Nonoriented discharges Obedience to mother Obedience to father Object orientation Orientation to humans Passiveness Reaction to father Reaction to mother Reaction to mother's punishment Unpleasant sensations
Touchette et al. (2013)	Included in meta-analysis	Quebec Newborn Twin Study	Canada	b134 Sleep functions	Daytime continuous sleep duration Night-time continuous sleep duration
Touwslager et al. (2011a)	Excluded from meta-analysis (duplicate data)	East Flanders Prospective Twin Survey	Belgium	b560 Growth maintenance functions	Growth in weight

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Touwslager et al. (2011b)	Included in meta-analysis	East Flanders Prospective Twin Survey	Belgium	b560 Growth maintenance functions	Weight gain
Tucker-Drob et al. (2011)	Included in meta-analysis	Early Childhood Longitudinal Study-Birth Cohort	United States	b163 Basic cognitive functions	Mental ability
van Dommelen et al. (2004)	Included in meta-analysis	Netherlands Twin Register	Netherlands	b560 Growth maintenance functions	Height Height deceleration Height jerk Height snap Height velocity Weight Weight deceleration Weight jerk Weight snap Weight velocity
Wang & Saudino (2012)	Included in meta-analysis	Boston University Twin Project	United States	b134 Sleep functions	Sleep problems
Whitfield et al. (2001)	Included in meta-analysis	Australian Twin Registry	Australia	b560 Growth maintenance functions	Weight
Wilson (1972)	Excluded from meta-analysis (duplicate data)	Louisville Twin Study	United States	b163 Basic cognitive functions	Cognitive ability
Wilson (1974)	Excluded from meta-analysis (duplicate data)	Louisville Twin Study	United States	b163 Basic cognitive functions	Cognitive ability

Table 4A (Continued)

Paper	Status	Study	Country	Category (using codes from the ICF-CY)	Phenotypes Extracted from Paper
Wilson (1978)	Excluded from meta-analysis (duplicate data)	Louisville Twin Study	United States	b163 Basic cognitive functions	Cognitive ability
Wilson (1983)	Excluded from meta-analysis (duplicate data)	Louisville Twin Study	United States	b163 Basic cognitive functions	Cognitive ability
Wilson (1984)	Included in meta-analysis	Louisville Twin Study	United States	b163 Basic cognitive functions b560 Growth maintenance functions	Cognitive ability Height
Wilson et al. (1972)	Included in meta-analysis	Louisville Twin Study	United States	b147 Psychomotor functions b163 Basic cognitive functions	Cognitive ability Motor development
Wilson & Matheny (1976)	Included in meta-analysis	Louisville Twin Study	United States	b163 Basic cognitive functions	Cognitive ability
Woodward et al. (2018)	Included in meta-analysis	Colorado Twin Registry: Longitudinal Twin Sample; Twin Infant Project	United States	d710 Basic interpersonal interactions	Child affection

Table 5A. Twin Studies Identified in the Systematic Literature Search, Presented Alphabetically by Study Name

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY[a])	Phenotypes (as labelled in included papers)	Papers
Australian Twin Registry	Australia	24	3808 (1799 MZ, 2009 DZ)	b560 Growth maintenance functions[b]	Weight	Whitfield et al. (2001)
Boston University Twin Project	United States	24	314 (145 MZ, 169 DZ)	b134 Sleep functions b140 Attention functions b147 Psychomotor functions b152 Emotional functions b164 Higher-level cognitive functions b560 Growth maintenance functions d710 Basic interpersonal interactions[c] d720 Complex interpersonal interactions	Activity Level (Home/Lab/Play), ADHD, Affective problems, Attention problems, Autistic-like traits, BMI, Callous Unemotional Traits, Difficult temperament, Externalizing, Inhibitory control, Non-social autistic-like traits, ODD, Sleep problems, Social autistic-like traits,	Flom & Saudino (2017) Flom & Saudino (2018) Flom et al. (2019) Gagne & Saudino (2010) Gagne et al. (2011) Gagne & Saudino (2016) Gagne et al. (2020) Ilott et al. (2010a) Ilott et al. (2010b) Micalizzi et al. (2016) Micalizzi et al. (2017) Ronald et al. (2010) Saudino (2012) Saudino & Zapfe (2008) Saudino et al. (2018) Silventoinen et al. (2016) Wang & Saudino (2012)
Brisbane Adolescent Twin Study	Australia	0	1330 (501 MZ, 829 DZ)	b560 Growth maintenance functions	BMI, Height, Weight,	Dubois et al. (2012) Hur et al. (2005)

---

[a] ICF-CY = International Classification of Functioning, Disability and Health, Children and Youth Version.[4] Definitions for each of the categories and subcategories can be found in the cited ICF-CY manual: https://apps.who.int/iris/handle/10665/43737.

[b] The prefix *b* is given to coded items within the ICF-CY component of Body Functions.

[c] The prefix *d* is given to items in the component of Activities and Participation.

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Child and Adolescent Twin Study in Sweden	Sweden	0	512 (172 MZ, 340 DZ)	b560 Growth maintenance functions	BMI, Height, Weight,	Dubois et al. (2012)
Childcare Centers of Tel Aviv and Haifa	Israel	0–12	163 (51 MZ, 112 DZ)	b147 Psychomotor functions b560 Growth maintenance functions	Age at transition to childhood Head circumference Head circumference growth Height Height growth Pulling up to a standing position Sitting up Turning over Walking five steps Weight Weight growth	German et al. (2015) Livshits et al. (2000) Peter et al. (1999)
Chinese National Twin Registry	China	0–36	3091 (1448 MZ, 1643 DZ)	b560 Growth maintenance functions	BMI Height Weight	Liu et al. (2015)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Collaborative Perinatal Project	United States	8	504 (189 MZ, 315 DZ)	b125 Dispositions and intra-personal functions b140 Attention functions b147 Psychomotor functions b163 Basic cognitive functions b560 Growth maintenance functions d710 Basic interpersonal interactions	Active manipulation Activity level Degree of social acceptance of examiner Degree of social contact with mother Interest in persons Interest in/responsiveness to people Mental development Physical development Pursuit persistence Response duration Speed of response Vigorous activity vs. psychomotor passivity	Goldsmith & Gottesman (1981) Nichols et al. (1974)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Colorado Twin Registry: Longitudinal Twin Sample	United States	14, 20, 24, 7–36	887 (494 MZ, 393 DZ)	b140 Attention functions b144 Memory functions b147 Psychomotor functions b152 Emotional functions b163 Basic cognitive functions b164 Higher-level cognitive functions b167 Mental functions of language d710 Basic interpersonal interactions	Activity level Affect Affect-extraversion Attention/persistence Behavioral inhibition Categorization Affection Disregard for others Emotionality Empathy Expressive language Frustration General cognitive ability Memory for Location Negative affect Negative hedonic tone Nonverbal Observed Disregard Overall mood Positive affect Positive hedonic tone Reactivity Receptive language Self-restraint Shyness Sociability Task orientation Verbal expressive Verbal receptive Visual attentiveness Word comprehension	Bishop et al. (2003b) Brant et al. (2009) Cherny et al. (1992) Cherny et al. (1994a) Cherny et al. (1994b) Dilalla et al. (1994) Emde et al. (1992) Friedman et al. (2011) Plomin et al. (1993) Reznick et al (1997) Rhee et al. (2007) Rhee et al. (2012) Rhee et al. (2013) Rhee et al. (2016) Robinson et al. (1992) Saudino et al. (1996) Schmitz et al. (1999) Smith et al. (2012) Woodward et al. (2018)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Colorado Twin Registry: Twin Infant Project	United States	7, 9, 7–36	168 (76 MZ, 92 DZ)	b140 Attention functions b152 Emotional functions b310 Voice functions d710 Basic interpersonal interactions	Affection for mother Child affection Enthusiasm for interaction with mother Negative affect Task orientation Watch mother Vocalize	DiLalla & Bishop (1996) Woodward et al. (2018)
Danish Twin Registry	Denmark	0	793 (141 MZ, 652 DZ)	b560 Growth maintenance functions	BMI Height Weight	Dubois et al. (2012)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Denver Twin Temperament Study	United States	3–16, 22	70 (35 MZ, 35 DZ)	b140 Attention functions b147 Psychomotor functions b152 Emotional functions d710 Basic interpersonal interactions d760 Family relationships	Activity level Approaching mother Approaching stranger Cuddliness with mother Cuddliness with stranger Difference of response between mother and stranger: approach Difference of response between mother and stranger: cuddliness Difference of response between mother and stranger: looking Difference of response between mother and stranger: positive vocalizations Difference of response between mother and stranger: proximity Difference of response between mother and stranger: quality of play Difference of response between mother and stranger: smiling Difference of response between mother and stranger: touches Distress to limitations Distress to novelty Duration of orienting Latency to approach stranger Looking at mother Looking at stranger Negative affect Positive affect Positive vocalization to mother Positive vocalization to stranger Proximity to mother Proximity to stranger Quality of play with mother Quality of play with stranger Resistance to soothing Separation distress Smiling and laughter Smiling at mother Smiling at stranger Touching mother Touching stranger	Goldsmith et al. (1999) Plomin & Rowe (1979)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Department of Psychology, University of Surrey	United Kingdom	0–24, 12–24	118 (41 MZ, 77 DZ)	b147 Psychomotor functions b152 Emotional functions b164 Higher-level cognitive functions d710 Basic interpersonal interactions	Activity level Emotionality Impulsivity Sociability	Stevenson & Fielding (1985)
Early Childhood Longitudinal Study-Birth Cohort	United States	9, 24	976 (238 MZ, 738 DZ)	b152 Emotional functions b163 Basic cognitive functions b167 Mental functions of language d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Attachment security Avoids others/not sociable Comfortable cuddly Cooperative Demanding/angry Dependency Enjoys company Expressive vocabulary Fussiness and demanding behavior Independent Mental ability Moody/unusual Positive and negative affect Seeks attention Upset by separation	Beaver et al. (2014) Jackson (2016) Roisman & Fraley (2006) Roisman & Fraley (2008) Tucker-Drob et al. (2011)
Early Reciprocal Social Behavior Study	United States	18, 24	317 (126 MZ, 191 DZ)	d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Behavior problems Competence Functional communication Reciprocal social behavior Social avoidance Social motivation Social orienting	Hawks & Marrus (2018) Marrus et al. (2015) Marrus et al. (2018) Marrus et al. (2020)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
East Flanders Prospective Twin Survey	Belgium	0–1, 1–6, 6–12, 12–24	280 (190 MZ, 90 DZ)	b560 Growth maintenance functions	Weight gain	Touwslager et al. (2011a) Touwslager et al. (2011b)
Gemini Study	United Kingdom	0, 3, 6, 16	2757 (1174 MZ, 1583 DZ)	b134 Sleep functions b147 Psychomotor functions b560 Growth maintenance functions d550 Eating	Activity level BMI Carbohydrate intake Daytime nap duration Emotional overeating Energy intake Enjoyment of food Fat intake First Crawl First Sit First Steps Food fussiness Food neophobia Food responsiveness Food weight Night awakenings Night-time sleep duration Protein intake Satiety responsiveness Size (weight) Slowness in eating Tempo (weight) Wake time Weight Weight change Weight velocity	Fisher et al. (2012) Herle et al. (2018) Johnson et al. (2011) Llewellyn et al. (2010) Pimpin et al. (2013) Silventoinen et al. (2016) Smith et al. (2017a) Smith et al. (2017b)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Guinea-Bissau Twin Study	Guinea-Bissau	0–2	108 (16 MZ, 92 DZ)	b560 Growth maintenance functions	BMI	Silventoinen et al. (2016)
Hungarian Twin Registry	Hungary	2	389 (230 MZ, 159 DZ)	b560 Growth maintenance functions	BMI	Silventoinen et al. (2016)
Italian Twin Registry (previously the Mercurio project)	Italy	12–24, 18, 24	7432 (3270 MZ, 4162 DZ)	b134 Sleep functions b560 Growth maintenance functions	BMI Cosleeping Diurnal sleep duration Night awakenings Nocturnal sleep duration Weight gain	Brescianini et al. (2011) Silventoinen et al. (2016)
Jackson Memorial Hospital/University of Miami Medical Center	United States	.5, 1, 3, 6, 9, 12	166 (67 MZ, 99 DZ)	b560 Growth maintenance functions	Length Weight	Levine et al. (1987)
Japanese Twin Cohort	Japan	1–2	2169 (1345 MZ, 824 DZ)	b560 Growth maintenance functions	BMI	Silventoinen et al. (2016)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Jumeaux et Plus ("Twins and more")	France	24	1950 (393 MZ, 1557 DZ)	b134 Sleep functions b140 Attention functions b147 Psychomotor functions b152 Emotional functions b279 Additional sensory functions d550 Eating d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Activity/Impulsivity Aggression/Defiance Attention Atypical Index Competence Compliance Depression withdrawal Dysregulation Eating problems Empathy Externalizing General anxiety Imitation/Play Inhibition to novelty Internalizing Maladaptive behaviors Mastery Motivation Negative affect Peer aggression Prosocial peer relations Sensory Sensitivity Separation distress Sleep problems Social relatedness	Saudino et al. (2008)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Louisville Twin Study	United States	0, 3, 6, 9, 12, 18, 24	615 (289 MZ, 326 DZ)	b140 Attention functions b147 Psychomotor functions b152 Emotional functions b163 Basic cognitive functions b279 Additional sensory functions b310 Voice functions b560 Growth maintenance functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Activity level Activity-awake Activity-sleep Affect-extraversion Approach/withdrawal Attachment Attention Auditory-visual Banging Behavioral inhibition Cognitive ability Cooperative Emotional tone Endurance Energy Extraversion Fearfulness Fine motor General temperament Goal directedness Gross motor Height Irritability Listening Looking Manipulating Motor development Mouthing: pacifier Mouthing: thumb Mouthing: toys Object orientation Primary cognition Reactivity Resistance to soothing Task orientation Tension Vocalize Weight	Davis et al. (2015) Finkel et al. (2000) Matheny (1980) Matheny (1983) Matheny (1984) Matheny (1989) Matheny et al. (1976) Riese (1990a) Riese (1990b) Wilson (1972) Wilson (1974) Wilson (1978) Wilson (1983) Wilson (1984) Wilson & Harpring (1972) Wilson & Matheny (1976)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Manitoba Twin Study	Canada	7	60 twin pairs (39 MZ, 21 DZ)	b 147 Psychomotor functions b 560 Growth maintenance functions	Activity level Head circumference Length Motor development Ponderal index Weight	Saudino & Eaton (1991)
Michigan State University Twin Registry	United States	24	10253 (3076 MZ, 7177 DZ)	b560 Growth maintenance functions	BMI	Silventoinen et al. (2016)
Minnesota Twin Family Study	United States	0	1068 (682 MZ, 386 DZ)	b560 Growth maintenance functions	Weight	Hur et al. (2005)
Mongolian Twin Registry	Mongolia	0–24	83 (36 MZ, 47 DZ)	b560 Growth maintenance functions	BMI	Silventoinen et al. (2016)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Moscow City twin sample	Russia	7–12	94 (49 MZ, 45 DZ)	b140 Attention functions b144 Memory functions b152 Emotional functions b164 Higher-level cognitive functions b279 Additional sensory functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions d760 Family relationships	Aggression towards father Aggression towards mother Autonomy Control Defensive reactions Dependence on mother EEG alpha frequency during darkness EEG mu gravity frequency during visual attention EEG spectral amplitude during darkness EEG spectral amplitude during visual attention Fear High tension Imitation of father Imitation of mother Love for father Love for mother Low tension Moderate tension Nonoriented discharges Obedience to mother Obedience to father Object orientation Orientation to humans Passiveness Reaction to father Reaction to mother Reaction to mother's punishment Unpleasant sensations Working memory	Orekhova et al. (2003) Pushina et al. (2005) Stroganova et al. (2000)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Multiple Births Foundation of Queen Charlotte's and Chelsea Hospital in London	United Kingdom	12–14	62 (30 MZ, 32 DZ)	b152 Emotional functions d760 Family relationships	Attachment disorganization Attachment security Temperamental reactivity	Bokhorst et al. (2003)
Netherlands Twin Register	The Netherlands	0, 1, 2, 3, 4, 6, 8, 9, 13, 14, 15.5, 24	16848 (5259 MZ, 11589 DZ)	b152 Emotional functions b560 Growth maintenance functions d760 Family relationships	Attachment disorganization Attachment security BMI Dependency (to father) Head circumference Height Height deceleration Height jerk Height snap Height velocity Infant-father attachment security Length Temperamental reactivity Weight Weight deceleration Weight jerk Weight snap Weight velocity	Bakermans-Kranenburg et al. (2004) Bokhorst et al. (2003) Boomsma et al. (1992) Mook-Kanamori et al. (2012) Silventoinen et al. (2016) Smit et al. (2010) van Dommelen et al. (2004)
Peri/Postnatal Epigenetic Twins Study	Australia	0–24	221 (91 MZ, 130 DZ)	b560 Growth maintenance functions	BMI	Silventoinen et al. (2016)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Pooled Sample of twins from Oregon, Washington, Colorado, Texas, and Wisconsin	United States	3–16	302 (121 MZ, 181 DZ)	b140 Attention functions b147 Psychomotor functions b152 Emotional functions	Activity level Distress to limitations Distress to novelty Duration of orienting Negative affect Positive affect Resistance to soothing Smiling and laughter	Goldsmith et al. (1999)
Project Metropolitan	Sweden	0	131 (28 MZ, 103 DZ)	b560 Growth maintenance functions	Height Weight	Akerman & Fischbein (1992)
Puerto Rican Infant Twin Study	United States	12, 0–32	865 (377 MZ, 488 DZ)	d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Difficult temperament Difficultness Inhibition Resistance to control Sociability Unadaptability Unsociability	Silberg et al. (2005)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Quebec Newborn Twin Study	Canada	0, 5, 6, 18, 19, 20, 18–24	1029 (419 MZ, 610 DZ)	b134 Sleep functions b152 Emotional functions b167 Mental functions of language b560 Growth maintenance functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions	BMI Cortisol reactivity Cortisol secretion in morning Cortisol secretion on awakening Daytime continuous sleep duration Difficult temperament Disregard for rules Disruptive behavior Emotional response to social stimuli Gaze aversion Height Motor activity during social stimuli Negative affect Nighttime continuous sleep duration Physical aggression Ratio of day/night sleep duration Self-contact during social stimuli Sleep terrors Social gaze Testosterone Vocabulary Weight	Caramaschi et al. (2012) Dionne et al. (2003b) Dionne et al. (2011) Dubois et al. (2007) Dubois et al. (2012) Forget-Dubois et al. (2007) Lacourse et al. 2014 Nguyen et al. (2008) Ouellet-Morin et al. (2008) Ouellet-Morin et al. (2009) Petitclerc et al. (2011) Schumann et al. (2017) Silventoinen et al. (2016) Soussignan et al. (2009) Touchette et al. (2013)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Registry of twins recruited in Japan from associations for parents of multiples.	Japan	0	1045 (775 MZ, 270 DZ)	b560 Growth maintenance functions	Weight	Hur et al. (2005)
Registry of twins who applied for the secondary school attached to the Faculty of Education at the University of Tokyo between 1981 and 2003.	Japan	0	1045 (775 MZ, 270 DZ)	b560 Growth maintenance functions	Weight	Hur et al. (2005)
Childcare Centers of Tel Aviv and Haifa	Israel	0–12	93 (64 DZ, 29 MZ)	b147 Psychomotor functions	Pulling up to a standing position Sitting up Turning over Walking five steps	Peter et al. (1999)
Seoul Twin Family Study	South Korea	0	686 (384 MZ, 302 DZ)	b560 Growth maintenance functions	Weight	Hur et al. (2005)
South Korea Twin Registry	South Korea	0	433 (255 MZ,178 DZ)	b560 Growth maintenance functions	Weight	Hur (2005)
Swedish Young Male Twins Study	Sweden	0, 12, 24	375 (231 MZ, 144 DZ)	b560 Growth maintenance functions	BMI Height	Silventoinen et al. (2007) Silventoinen et al. (2008) Silventoinen et al. (2016)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Taipei City Teaching Hospitals Twin Study	Taiwan	1, 2, 4, 6, 9, 12	521 (428 MZ, 93 DZ)	b125 Dispositions and intra-personal functions b140 Attention functions b147 Psychomotor functions b152 Emotional functions b560 Growth maintenance functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions	Activity level Adaptability Approach/withdrawal Arm circumference Attention/persistence Chest circumference Distractibility Head circumference Height Intensity of reaction Quality of mood Rhythmicity Threshold of responsiveness Weight	Chen et al. (1990a) Chen et al. (1990b)
Tokyo Twin Cohort Project	Japan	0, 4, 10, 19	1728 (669 MZ, 1059 DZ)	b134 Sleep functions b560 Growth maintenance functions d550 Eating d710 Basic interpersonal interactions	Chest circumference Head circumference Head circumference growth Height Milk consumption Mimic, point gazing, joint attention Rhythmicity Sociocognitive abilities Time to fall asleep Weight	Ando et al. (2006) Fujisawa et al. (2012)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
Twin study in The Gambia	The Gambia	0, 5–18	84 (22 MZ, 62 DZ)	b147 Psychomotor functions b560 Growth maintenance functions	Crawl Length Maintain head Roll over Sitting Sitting without support Stand holding on something Take two steps Walk holding on something Weight	Goetghebuer et al. (2003)
Twin Study of Child and Adolescent Development	Sweden	0	823 (376 MZ, 447 DZ)	b560 Growth maintenance functions	BMI Height Weight	Dubois et al. (2012)
Twins born in two South Korean hospitals (1998–2003)	South Korea	0	603 (338 MZ, 265 DZ)	b560 Growth maintenance functions	Weight	Hur et al. (2005)
Twins Early Development Study	United Kingdom	24	9065 (3082 MZ, 5983 DZ)	b140 Attention functions b147 Psychomotor functions b163 Basic cognitive functions b167 Mental functions of language b560 Growth maintenance functions d710 Basic interpersonal interactions d720 Complex interpersonal interactions	ADHD Behavior problems BMI General cognitive ability Grammar Hyperactivity Nonverbal cognitive ability Prosocial behavior Verbal ability	Dale et al. (2000) Dionne et al. (2003a) Galsworthy et al. (2000) Knafo & Plomin (2006) Koeppen-Schomerus et al. (2003) Kuntsi et al. (2005) Price et al. (2000) Price et al. (2005) Silventoinen et al. (2016) Spinath et al. (2003)

Table 5A (Continued)

Study name	Country	Age (months)	n twin pairs	Category (using codes from the ICF-CY)	Phenotypes (as labelled in included papers)	Papers
UNC Early Brain Development Study	United States	0, 0–3	180 (63 MZ, 117 DZ)	s110 Structure of brain[d]	Cerebellum Corpus callosum Cortical grey matter Cortical surface area Cortical thickness Cortical unmyelinated white matter Frontal grey matter Frontal unmyelinated white matter Intracranial volume Lateral ventricles Left hemisphere grey matter Left hemisphere total Left hemisphere unmyelinated white matter Occipital grey matter Occipital unmyelinated white matter Parietal grey matter Parietal unmyelinated white matter Prefrontal grey matter Prefrontal unmyelinated white matter Right hemisphere grey matter Right hemisphere total Right hemisphere unmyelinated white matter Subcortical grey matter Total cerebrospinal fluid Total early myelinated white matter Total frontal Total grey matter Total occipital Total parietal Total prefrontal Total unmyelinated white matter	Gilmore et al. (2010) Jha et al. (2018)

---

[d] The prefix *s* is given to items in the component of Body Structures.

Table 5A (Continued)

University of Sao Paulo Longitudinal Twin Study	Brazil	0–6	17 (10 MZ, 7 DZ)	b134 Sleep functions	Emergence of cortisol circadian rhythm	Custodio et al. (2007)
Waisman Center Birth to 3 year project	United States	6, 12	531 (180 MZ, 351 DZ)	b152 Emotional functions	Positive affect Smiling and laughter	Planalp et al. (2017)
West Japan Twins and Higher Order Multiple Births Registry	Japan	0, 1–3, 3–5, 5–7, 7–9, 9–11, 11–13, 12, 24	767 (407 MZ, 360 DZ)	b560 Growth maintenance functions	BMI Chest circumference Chest circumference increase Head circumference Height	Silventoinen et al. (2011a) Silventoinen et al. (2011b) Silventoinen et al. (2012) Silventoinen et al. (2016)
Western Australian Twin Registry	Australia	24	473 (160 MZ, 313 DZ)	b167 Mental functions of language	Combining words Late language acquisition Use of finiteness grammatical markers Words Produced	Rice et al. (2014)
Wisconsin Twin Panel	United States	12	735 (261 MZ, 474 DZ)	b152 Emotional functions	Anger Distress to limitations	Gagne & Goldsmith (2011)

Table 6A. Phenotypes Identified in the Systematic Literature Search Coded Using the Classification System from the ICF-CY

ICF-CY[a] component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Global mental functions	b125 Dispositions and intra-personal functions	b1251 Responsivity	• Response duration • Speed of response • Threshold of responsiveness
Body functions	Mental functions	Global mental functions	b134 Sleep functions	b1340 Amount of sleep	• Cortisol secretion in morning • Cortisol secretion on awakening • Daytime continuous sleep duration • Daytime nap duration • Diurnal sleep duration • Emergence of the cortisol circadian rhythm • Night-time continuous sleep duration • Night-time sleep duration • Nocturnal sleep duration • Rhythmicity • Wake time
Body functions	Mental functions	Global mental functions	b134 Sleep functions	b1341 Onset of sleep	• Time to fall asleep
Body functions	Mental functions	Global mental functions	b134 Sleep functions	b1342 Maintenance of sleep	• Night awakenings
Body functions	Mental functions	Global mental functions	b134 Sleep functions	b1343 Quality of sleep	• Sleep problems • Sleep terrors
Body functions	Mental functions	Global mental functions	b134 Sleep functions	b1348 Sleep functions, other specified	• Ratio of day/night sleep duration

---

[a] ICF-CY = International Classification of Functioning, Disability and Health, Children and Youth Version.[4] Definitions for each of the categories and subcategories can be found in the cited ICF-CY manual: https://apps.who.int/iris/handle/10665/43737.

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Specific mental functions	b140 Attention functions	b1400 Sustaining attention	• ADHD • Attention • Attention problems • Attention/persistence • Distractibility • Duration of orienting • Endurance • Goal directedness • Listening • Looking • Mastery motivation • Object orientation • Pursuit persistence • Task orientation • Visual attentiveness
Body functions	Mental functions	Specific mental functions	b140 Attention functions	b1408 Attention functions, other specified	• Mu frequency during visual attention • Spectral amplitude during visual attention
Body functions	Mental functions	Specific mental functions	b144 Memory functions	b1440 Short-term memory	• Memory for location • Working memory

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Specific mental functions	b147 Psychomotor functions	b1470 Psychomotor control	• Activity • Activity level • Activity level: Home • Activity level: Lab • Activity level: Play • Activity-awake • Activity-sleep • Activity/impulsivity • Energy • Hyperactivity • Vigorous activity vs. psychomotor passivity

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Specific mental functions	b147 Psychomotor functions	b1472 Organization of psychomotor functions	• Active manipulation • Banging • Crawl • Fine motor • First crawl • First sit • First steps • Gross motor • Maintain head • Manipulating • Motor development • Mouthing: pacifier • Mouthing: thumb • Mouthing: toys • Pulling up to standing position • Roll over • Sitting • Sitting up • Sitting without support • Stand holding on something • Take two steps • Turning over • Walk holding on something • Walking five steps

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Specific mental functions	b152 Emotional functions	b1521 Regulation of emotion	• Affective problems • Cortisol reactivity • Distress to limitations • Emotionality • Moody/unusual • Quality of mood • Reactivity • Resistance to soothing • Temperamental reactivity
Body functions	Mental functions	Specific mental functions	b152 Emotional functions	b1522 Range of emotion	• Affect • Anger • Demanding/angry • Depression withdrawal • Emotional tone • Fear • Fearfulness • Frustration • Fussiness and demanding behavior • General anxiety • High tension • Intensity of reaction • Internalizing • Irritability • Low tension • Moderate tension • Negative affect • Negative hedonic tone • Nonoriented discharges • Overall mood • Positive Affect • Positive and negative affect • Positive hedonic tone • Smiling and laughter

					- Tension

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Specific mental functions	b163 Basic cognitive functions	N/A	• Cognitive ability • General cognitive ability • Mental ability • Mental development • Nonverbal • Nonverbal cognitive ability • Nonverbal cognitive development • Primary cognition
Body functions	Mental functions	Specific mental functions	b164 Higher-level cognitive functions	b1641 Higher-level cognitive functions, other specified	• Categorization • Control • Impulsivity • Inhibitory control • Self-restraint

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Mental functions	Specific mental functions	b167 Mental functions of language	b1670 Reception of language	• Reception of language • Verbal receptive • Word comprehension
Body functions	Mental functions	Specific mental functions	b167 Mental functions of language	b1671 Expression of language	• Combining words • Expressive language • Expressive vocabulary • Grammar • Late language acquisition • Use of finiteness grammatical markers • Verbal ability • Verbal expressive • Vocabulary • Words Produced
Body functions	Sensory functions and pain	Additional sensory functions	b279 Additional sensory functions, other specified and unspecified	N/A	• Auditory-visual • Sensory Sensitivity • Unpleasant sensations
Body functions	Voice and speech functions	N/A	b310 Voice functions	b3100 Production of voice	• Vocalize

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body functions	Functions of the digestive, metabolic and endocrine systems	Functions related to the metabolism and endocrine system	b560 Growth maintenance functions	N/A	• Age at transition to childhood • Arm circumference • BMI • Chest circumference • Chest circumference increase • Head circumference • Head circumference growth • Head circumference growth curve parameters • Height • Height growth curve parameters • Height deceleration • Height jerk • Height snap • Height velocity • Length • Physical development • Ponderal index • Weight • Weight growth • Weight growth curve parameters • Weight deceleration • Weight gain • Weight jerk • Weight size • Weight snap • Weight tempo • Weight velocity

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Activities and participation	Self-care	N/A	d550 Eating	N/A	• Carbohydrate intake • Eating problems • Emotional overeating • Energy intake • Enjoyment of food • Fat intake • Food fussiness • Food neophobia • Food responsiveness • Food weight • Milk consumption • Protein intake • Satiety responsiveness • Slowness in eating
Activities and participation	Interpersonal interactions and relationships	General interpersonal interactions	d710 Basic interpersonal interactions	d7100 Respect and warmth in relationships	• Affection for mother • Callous unemotional behavior • Callous unemotional traits • Child affection • Disregard for others • Empathy • Observed Disregard • Prosocial behavior • Prosocial Peer Relations

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Activities and participation	Interpersonal interactions and relationships	General interpersonal interactions	d710 Basic interpersonal interactions	d7104 Social cues in relationships	• Affect-extraversion • Approach/withdrawal • Approaching mother • Approaching stranger • Avoids others/not sociable • Behavioral inhibition • Competence • Degree of social acceptance of examiner • Degree of social contact with mother • Emotional response to social stimuli • Enjoys company • Enthusiasm for interaction with mother • Extraversion • Gaze aversion • Imitation/Play • Imitation of father • Imitation of mother • Inhibition • Inhibition to novelty • Interest in persons • Interest in/responsiveness to people • Latency to approach stranger • Looking at mother • Looking at stranger • Mimic, point gazing, joint attention • Orientation to humans • Positive vocalization to mother • Positive vocalization to stranger • Proximity to stranger • Quality of play with mother • Quality of play with stranger • Reaction to father • Reaction to mother • Reaction to mother's punishment • Reciprocal social behavior • Seeks attention • Shyness • Smiling at mother • Smiling at stranger • Sociability • Social autistic-like traits • Social gaze • Social relatedness • Sociocognitive abilities • Unsociability • Watch Mother

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Activities and participation	Interpersonal interactions and relationships	General interpersonal interactions	d710 Basic interpersonal interactions	d7105 Physical contact in relationships	• Comfortable cuddly • Cuddliness with mother • Cuddliness with stranger • Touching mother • Touching stranger
Activities and participation	Interpersonal interactions and relationships	General interpersonal interactions	d710 Basic interpersonal interactions	d7106 Differentiation of familiar persons	• Difference of response between mother and stranger: approach • Difference of response between mother and stranger: cuddliness • Difference of response between mother and stranger: looking • Difference of response between mother and stranger: positive vocalizations • Difference of response between mother and stranger: proximity • Difference of response between mother and stranger: quality of play • Difference of response between mother and stranger: smiling • Difference of response between mother and stranger: touches

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Activities and participation	Interpersonal interactions and relationships	General interpersonal interactions	d720 Complex interpersonal interactions	d7202 Regulating behaviors within interactions	• Adaptability • Aggression towards father • Aggression towards mother • Aggression/defiance • Defensive reactions • Difficult temperament • Difficultness • Disruptive behavior • Externalizing • Peer aggression • Physical aggression • Unadaptability
Activities and participation	Interpersonal interactions and relationships	General interpersonal interactions	d720 Complex interpersonal interactions	d7203 Interacting according to social rules	• Behavior problems • Compliance • Cooperative • Disregard for rules • Maladaptive behavior • Obedience to mother • Obedience to father • ODD • Resistance to control

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Activities and participation	Interpersonal interactions and relationships	Particular interpersonal relationships	d760 Family relationships	d7601 Child-parent relationships	• Attachment • Attachment disorganization • Attachment security • Dependence on mother • Dependency • Dependency to father • Independent • Infant-father attachment security • Proximity to mother • Separation distress • Temperamental dependency • Upset by separation
Body structures	Structures of the nervous system	N/A	s110 Structure of brain	s1100 Structure of cortical lobes	• Frontal grey matter • Occipital grey matter • Parietal grey matter • Prefrontal grey matter • Total Frontal • Total Occipital • Total Parietal • Total Prefrontal
Body structures	Structures of the nervous system	N/A	s110 Structure of brain	s1104 Structure of cerebellum	• Cerebellum

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Body structures	Structures of the nervous system	N/A	s110 Structure of brain	s1107 Structure of white matter	• Total unmyelinated white matter • Total early myelinated white matter • Cortical unmyelinated white matter • Prefrontal unmyelinated white matter • Frontal unmyelinated white matter • Parietal unmyelinated white matter • Occipital unmyelinated white matter • Right hemisphere unmyelinated white matter • Left hemisphere unmyelinated white matter • Corpus Callosum
Body structures	Structures of the nervous system	N/A	s110 Structure of brain	s1108 Structure of brain, other specified	• Cortical grey matter • Cortical surface area • Cortical thickness • Intracranial volume • Lateral Ventricles • Left hemisphere grey matter • Left hemisphere total • Right hemisphere grey matter • Right hemisphere total • Subcortical grey matter • Total cerebrospinal fluid • Total grey matter

Table 6A (Continued)

ICF-CY component	ICF-CY domain	ICF-CY first level item	ICF-CY second level item (category)	ICF-CY third-level item (sub-category)	Phenotypes (as labelled in included papers)
Uncategorized	N/A	N/A	N/A	N/A	• Autistic-like traits • Alpha frequency during darkness • Atypical index • Autonomy • Cosleeping • Distress to novelty • Dysregulation • General temperament • Love for father • Love for mother • Motor activity during social stimuli • Non-social autistic-like traits • Passiveness • Restrictive repetitive behavior • Rhythmicity • Self-contact/comfort during social stimuli • Spectral amplitude during darkness • Testosterone

Table 7A. Estimates and Cohorts in Phenotypic Categories by Rater

ICF-CY category	Rater	k_{cohort}	$k_{estimate}$	$\%_{estimate}$
b134 Sleep	Parent	6	43	87.76
	Observer	0	0	0.00
	Other	2	6	12.24
b140 Attention	Parent	6	33	18.86
	Observer	4	66	37.71
	Other	1	76	43.43
b147 Psychomotor	Parent	10	45	29.80
	Observer	5	84	55.63
	Other	3	22	14.57
b152 Emotional	Parent	11	90	41.67
	Observer	6	114	52.78
	Other	2	12	5.56
b163 Basic cognitive	Parent	1	7	14.89
	Observer	4	40	85.11
	Other	0	0	0.00
b167 Language	Parent	5	38	39.58
	Observer	1	58	60.42
	Other	0	0	0.00
b560 Growth	Parent	0	0	0.00
	Observer	0	0	0.00
	Other	24	465	100.00
d710 Basic interpersonal	Parent	11	136	38.20
	Observer	7	202	56.74
	Other	2	18	5.06
d720 Complex interpersonal	Parent	8	59	80.82
	Observer	1	4	5.48
	Other	1	10	13.70
d760 Family relationships	Parent	3	17	58.62
	Observer	3	10	34.48
	Other	1	2	6.90
Total	Parent	22	468	28.24
	Observer	12	578	34.88
	Other	27	611	36.87

Note. k_{cohort} = number of independent twin cohorts. $k_{estimate}$ = number of estimates (twin correlations).

Table 8A. Multilevel Random Effects Models of Phenotypic Sub-Categories

ICF-CY sub-category	k_{cohort}	$k_{estimate}$	n_{MZ}	n_{DZ}	r_{MZ} pooled	r_{MZ} 95% CI	r_{DZ} pooled	r_{DZ} 95% CI	h^2 pooled	h^2 95% CI	h^2 p	c^2 pooled	c^2 95% CI	c^2 p	e^2 pooled	e^2 95% CI	e^2 p	I^2 Level 1	I^2 Level 2	I^2 Level 3
Sustaining attention	10	101	3007	6135	.62	.49–.75	.37	.24–.50	.50	.14–.75	.007	.12	0–.41	.424	.38	.25–.51	<.001	6.53	42.17	51.30
Psychomotor control	11	83	3059	6287	.62	.48–.76	.29	.15–.43	.61	.26–.75	.001	.00	0–.27	1.000	.39	.26–.51	<.001	3.93	18.92	77.15
Organization of psychomotor	6	68	1022	1620	.79	.70–.88	.54	.45–.64	.49	.22–.76	<.001	.29	.08–.51	.006	.21	.12–.30	<.001	1.77	43.00	55.23
Regulation of emotion	10	48	980	1692	.53	.41–.65	.33	.20–.45	.41	.07–.66	.016	.12	0–.39	.379	.47	.35–.59	<.001	16.31	13.02	70.67
Range of emotion	11	168	1514	3271	.59	.49–70	.38	.28–.49	.42	.13–.69	.005	.17	0–.41	.153	.41	.30–.51	<.001	5.19	71.82	23.00
Expression of language	5	46	2244	2864	.86	.75–.98	.73	.61–.84	.27	0–.59	.099	.59	.34–.85	<.001	.14	.03–.25	<.001	0.21	58.26	41.53
Respect and warmth	7	39	3078	6131	.63	.45–.80	.53	.35–.70	.20	0–.69	.405	.42	.04–.69	.028	.37	.20–.53	.015	0.52	81.71	17.78
Social cues	15	285	1983	4030	.58	.46–.70	.38	.26–.50	.41	.06–.70	.021	.18	0–.46	.207	.42	.30–.54	<.001	2.20	20.69	77.10
Regulating behaviors	6	43	1056	2385	.72	.58–.87	.43	.28–.58	.58	.18–.87	.004	.14	0–.47	.393	.28	.14–.42	<.001	2.47	36.86	60.67
Social rules	9	30	3173	5070	.72	.60–.83	.56	.45–.68	.31	0–.63	.052	.41	.16–.66	.001	.28	.18–.39	<.001	1.93	0.00	98.07

Note. k_{cohort} = number of independent twin cohorts. $k_{estimate}$ = number of estimates (twin correlations). n_{MZ} = number of monozygotic (MZ) twin pairs. n_{DZ} = number of dizygotic (DZ) twin pairs. r_{MZ} = MZ twin correlation. $_{95\%\ CI}$ = 95% confidence interval. r_{DZ} = DZ twin correlation. h^2 = heritability. p = p-value. c^2 = shared environment. e^2 = nonshared environment. I^2 = heterogeneity. $_{\text{Level 1}}$ = sampling variance. $_{\text{Level 2}}$ = within-cohort variance in outcome measurement. $_{\text{Level 3}}$ = between-cohort variance.

Table 9A. Multilevel Random Effects Models of Parent and Observer Ratings of Three Phenotypic Categories

ICF-CY category and rater	k_{cohort}	$k_{estimate}$	n_{MZ}	n_{DZ}	r_{MZ} pooled	r_{MZ} 95% CI	r_{DZ} pooled	r_{DZ} 95% CI	h^2 pooled	h^2 95% CI	h^2 p	c^2 pooled	c^2 95% CI	c^2p	e^2 pooled	$e^2{}_{95\% \ CI}$	e^2p	I^2 Level 1	I^2 Level 2	I^2 Level 3
Psychomotor (obs)	5	84	503	588	0.55	.37–.74	0.38	.20–.56	0.35	0–.74	.182	0.21	0–.57	.313	0.45	.27–.63	<.001	7.84	45.02	47.14
Psychomotor (par)	10	45	2862	5732	0.69	.52–.85	0.30	.13–.47	0.67	.30–.82	.001	0.00	0–.29	1.000	0.33	.19–.47	<.001	0.88	2.92	96.20
Emotional (obs)	6	114	793	1261	0.40	.27–.53	0.26	.13–.39	0.28	0–.53	.129	0.12	0–.40	.391	0.60	.47–.73	<.001	16.49	43.14	40.38
Emotional (par)	11	90	1579	3492	0.67	.60–.74	0.42	.35–.49	0.49	.28–.71	<.001	0.17	.01–.35	.033	0.33	.26–.40	<.001	5.69	94.31	0.00
Basic interpersonal (obs)	7	202	839	945	0.37	.25–.49	0.27	.15–.39	0.20	0–.50	.254	0.17	0–.40	.221	0.63	.51–.75	<.001	18.76	32.15	49.09
Basic interpersonal (par)	11	136	3518	7229	0.70	.57–.82	0.42	.29–.55	0.55	.19–.83	.002	0.15	0–.44	.307	0.30	.18–.43	<.001	1.21	32.18	66.61

Note. k_{cohort} = number of independent twin cohorts. $k_{estimate}$ = number of estimates (twin correlations). n_{MZ} = number of monozygotic (MZ) twin pairs. n_{DZ} = number of dizygotic (DZ) twin pairs. r_{MZ} = MZ twin correlation. $_{95\% \ CI}$ = 95% confidence interval. r_{DZ} = DZ twin correlation. h^2 = heritability. p = p-value. c^2 = shared environment. e^2 = nonshared environment. I^2 = heterogeneity. $_{\text{Level 1}}$ = sampling variance. $_{\text{Level 2}}$ = within-cohort variance in outcome measurement. $_{\text{Level 3}}$ = between-cohort variance. obs = observer rated. par = parent rated.

Table 10A. Tests for Publication Bias on Twin Correlations by Phenotype Category

	Egger's r_{MZ}			Egger's r_{DZ}		
ICF-CY category	$k_{estimate}$	z	p	$k_{estimate}$	z	p
b134 Sleep	24	-10.38	< .001	25	-3.55	< .001
b140 Attention	86	-8.51	< .001	89	-1.88	.061
b147 Psychomotor	75	-11.20	< .001	76	-4.17	< .001
b152 Emotional	105	-10.68	< .001	111	-4.76	< .001
b163 Basic cognitive	23	-6.26	< .001	24	-7.24	< .001
b167 Language	48	-22.50	< .001	48	-15.36	< .001
b560 Growth	216	-7.77	< .001	249	-4.78	< .001
d710 Basic interpersonal	174	-11.55	< .001	182	-10.05	< .001
d720 Complex interpersonal	34	-7.07	< .001	39	-3.40	< .001
d760 Family relationships	14	-4.84	< .001	15	-1.86	.062

Note. Egger's rMZ = Egger's test on monozygotic twin correlations. Egger's rMZ = Egger's test on dizygotic twin correlations. $k_{estimate}$ = number of estimates (twin correlations).

Table 11A. Tests for Publication Bias on Phenotype Categories by Variance Component in Phenotype Categories With ≥ 10 Estimates

		Egger's h^2		Egger's c^2		Egger's e^2	
ICF-CY category	$k_{estimate}$	*z*	*p*	*z*	*p*	*z*	*p*
b134 Sleep	20	-1.80	.072	-3.62	< .001	-0.78	.438
b140 Attention	12	-1.12	.264	-0.84	.402	-2.10	.036
b147 Psychomotor	15	-5.26	< .001	-3.73	< .001	0.68	.494
b152 Emotional	20	-3.07	.002	-2.67	.008	-1.04	.297
b560 Growth	90	-0.87	.385	-5.64	< .001	-1.07	.285
d710 Basic interpersonal	41	-2.47	.013	-8.40	< .001	-2.21	.027
d720 Complex interpersonal	15	-3.07	.002	-5.52	< .001	-1.65	.099

Note. Because many studies only reported twin correlations, and not h^2, c^2 and e^2 estimates, the number of estimates included in the Egger's tests of h^2, c^2 and e^2 was smaller for all phenotypic categories than the number of estimates included in Egger's tests of twin correlations. There were too few estimates (< 10 estimates) to meet the study criteria to create funnel plots or run Egger's tests on estimates of h^2, c^2 and e^2 for 'basic cognitive functions', 'mental functions of language' and 'family relationships'. Egger's h^2 = Egger's test on heritability estimates. Egger's c^2 = Egger's test on shared environment estimates. Egger's e^2 = Egger's test on nonshared environment estimates. $k_{estimate}$ = number of estimates (twin correlations).

Figure 1A. Prisma Flow Diagram

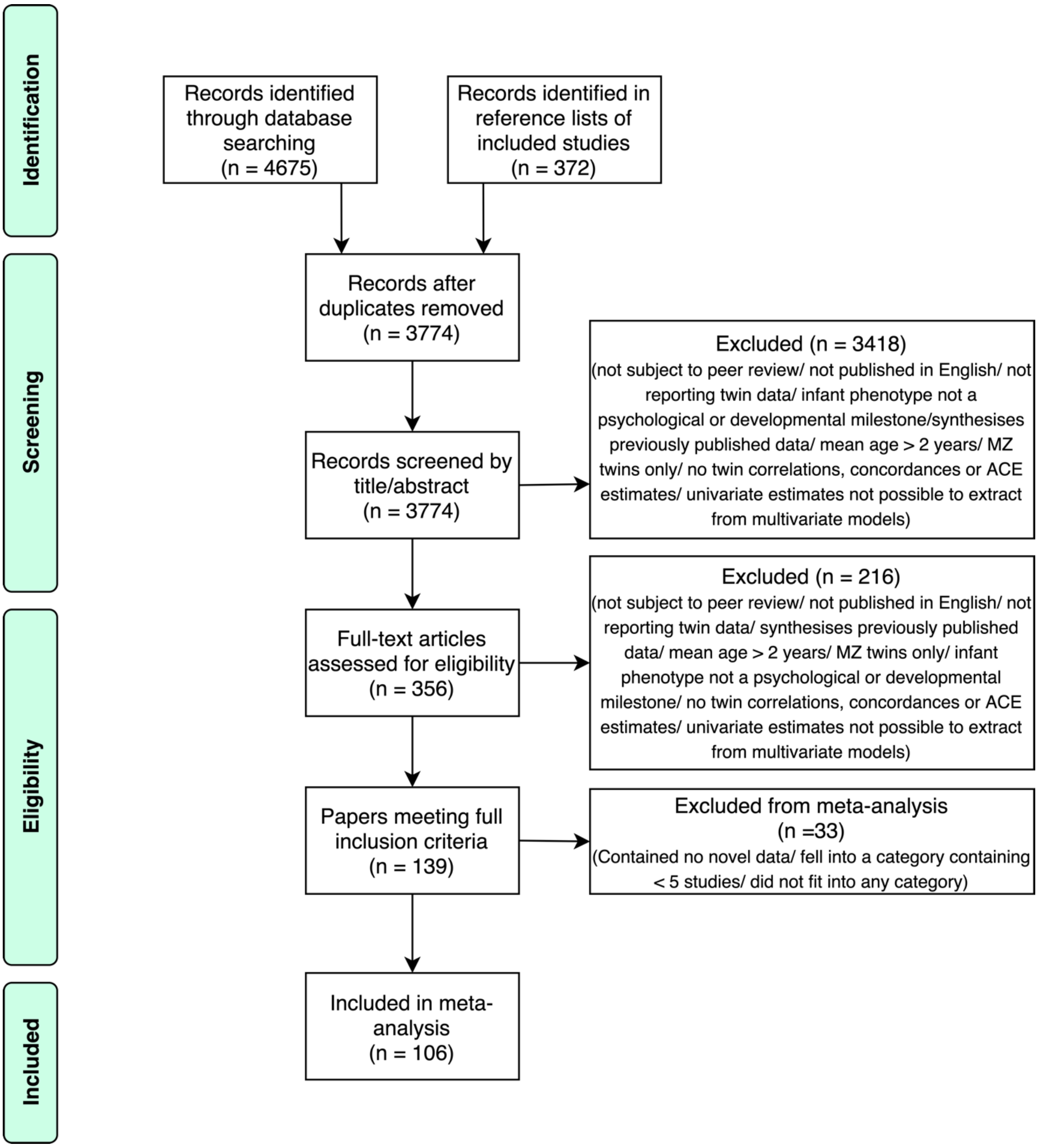
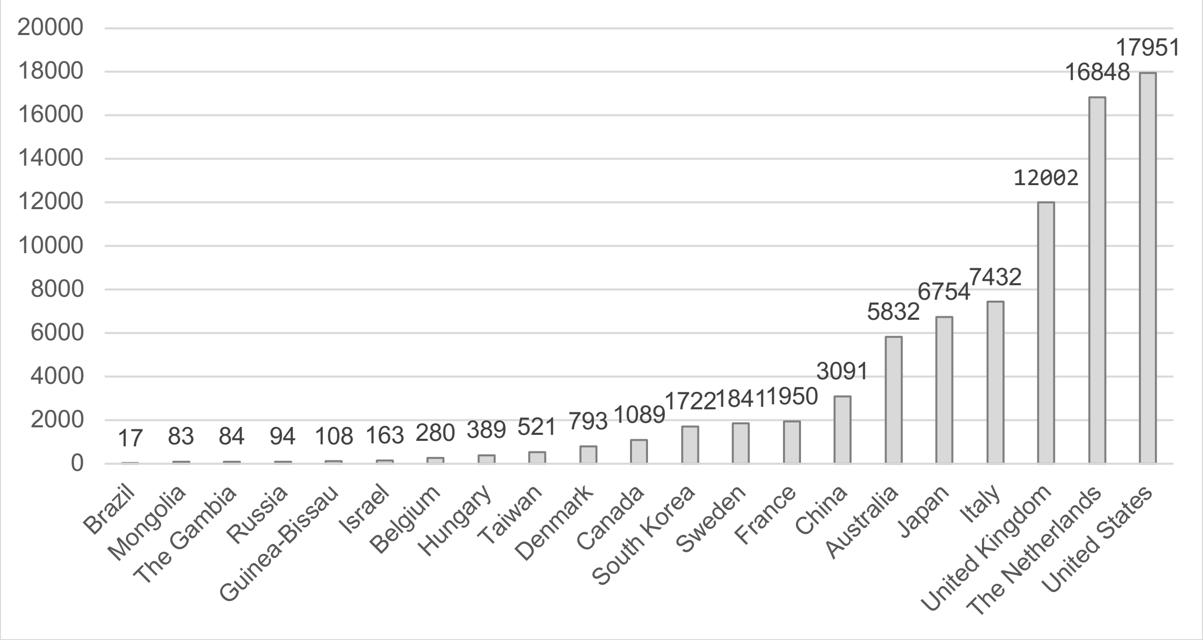
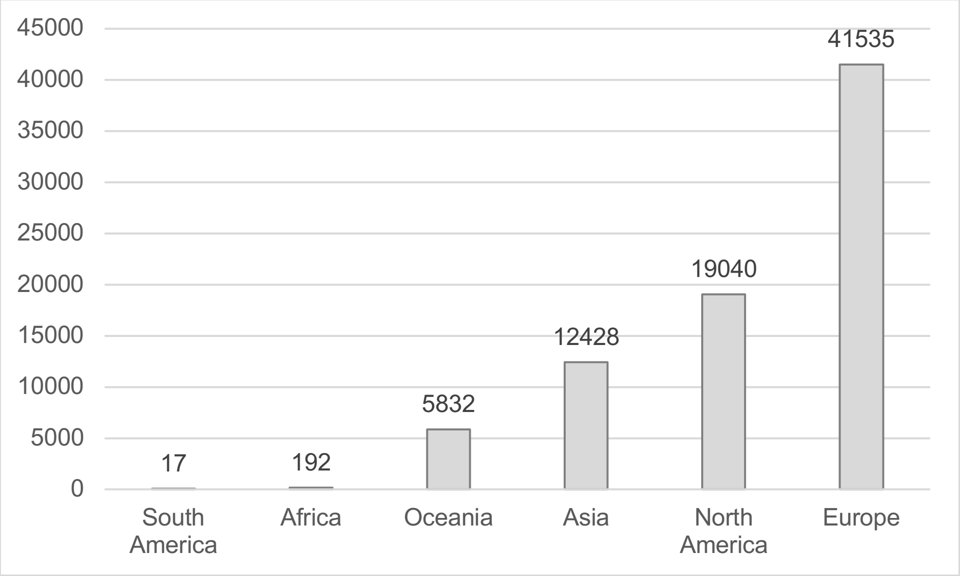


Figure 2A. Bar Chart of Number of Twin Pairs by Country



Note. Total number of twin pairs = 79,044

Figure 3A. Bar Chart of Number of Twin Pairs by Continent



Note. Total number of twin pairs = 79,044

Figure 4A. Sleep Functions Forest Plot

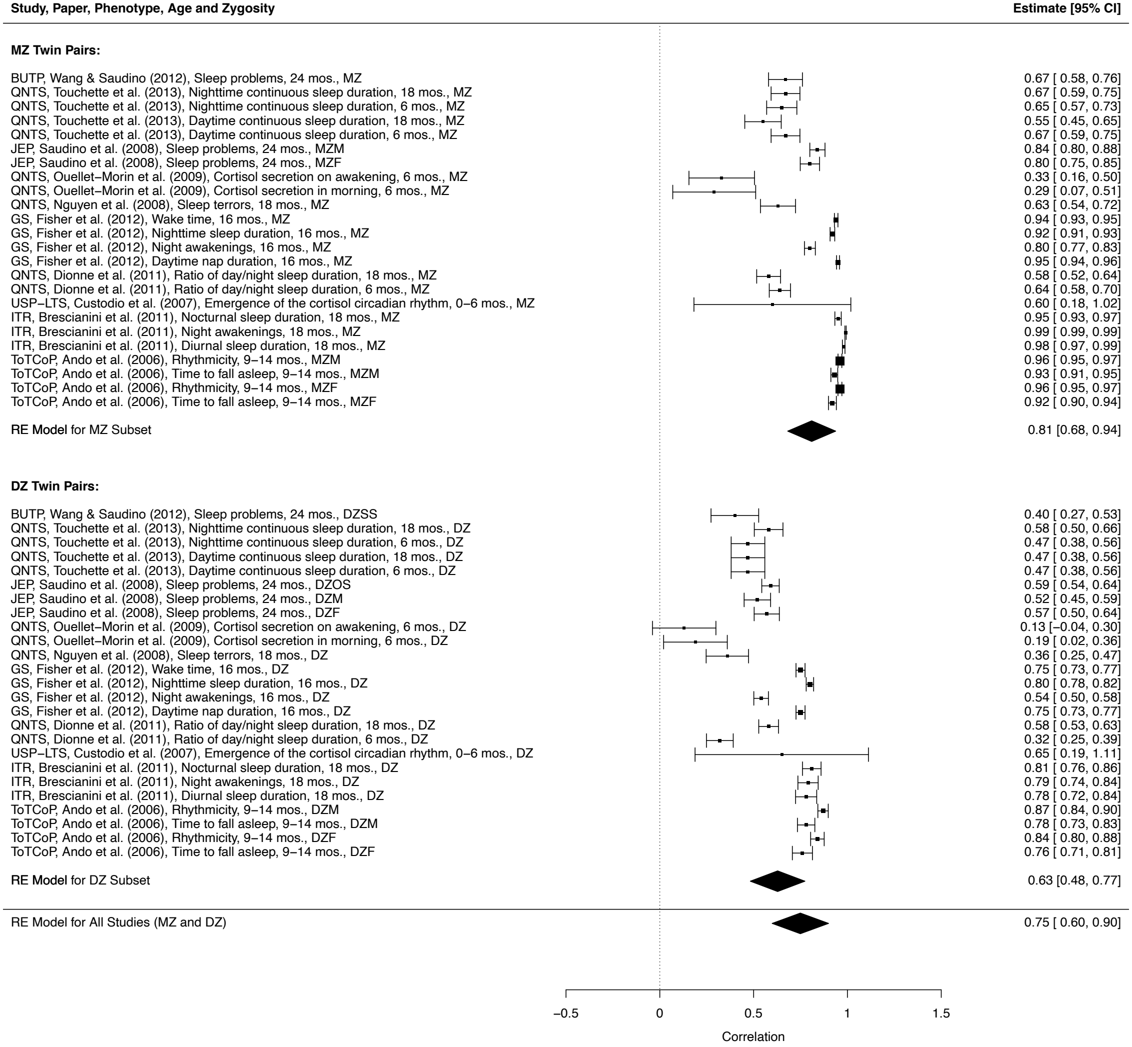
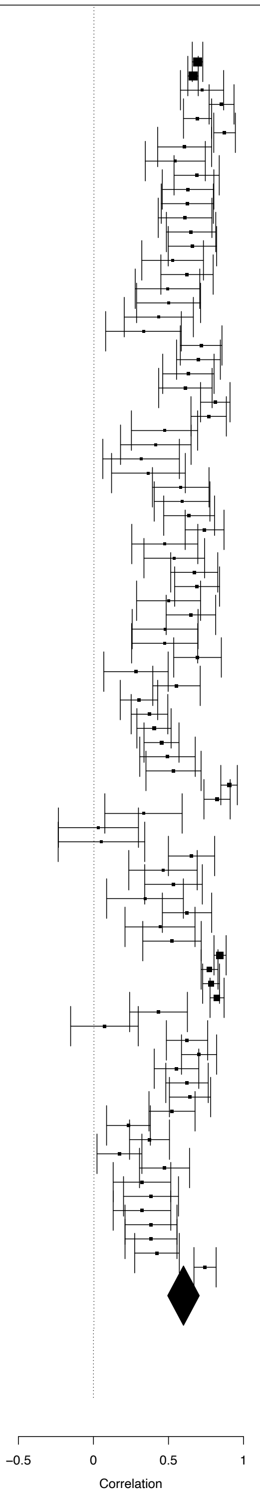


Figure 5A. Attention Functions Forest Plot



Study, Paper, Phenotype, Age and Zygosity	**Estimate [95% CI]**
MZ Twin Pairs:	
TEDS, Price 2005, ADHD, 24 mos., MZM	0.69 [0.66, 0.72]
TEDS, Price 2005, ADHD, 24 mos., MZF	0.66 [0.63, 0.69]
TCTH, Chen 1990, Distractibility, 6 mos., MZ	0.72 [0.58, 0.86]
TCTH, Chen 1990, Attention/persistence, 6 mos., MZ	0.85 [0.77, 0.93]
Pooled_sample, Goldsmith 1999, Duration of orienting, 3–16 mos., MZ	0.69 [0.60, 0.78]
MCTS, Stroganova 2000, Object orientation, 7–12 mos., MZ	0.87 [0.80, 0.94]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta T6, 8–11 mos., MZ	0.60 [0.42, 0.78]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta T5, 8–11 mos., MZ	0.54 [0.34, 0.74]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta PO4, 8–11 mos., MZ	0.68 [0.53, 0.83]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta PO3, 8–11 mos., MZ	0.63 [0.46, 0.80]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta O2, 8–11 mos., MZ	0.62 [0.45, 0.80]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta O1, 8–11 mos., MZ	0.61 [0.43, 0.79]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta FC4, 8–11 mos., MZ	0.65 [0.48, 0.81]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta FC3, 8–11 mos., MZ	0.66 [0.49, 0.82]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta F8, 8–11 mos., MZ	0.52 [0.32, 0.73]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta F7, 8–11 mos., MZ	0.62 [0.45, 0.79]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta AF4, 8–11 mos., MZ	0.49 [0.28, 0.70]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Theta AF3, 8–11 mos., MZ	0.50 [0.28, 0.71]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta T6, 8–11 mos., MZ	0.43 [0.20, 0.66]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta T5, 8–11 mos., MZ	0.33 [0.08, 0.58]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta PO4, 8–11 mos., MZ	0.71 [0.58, 0.85]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta PO3, 8–11 mos., MZ	0.70 [0.55, 0.84]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta O2, 8–11 mos., MZ	0.63 [0.46, 0.80]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta O1, 8–11 mos., MZ	0.61 [0.43, 0.79]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta FC4, 8–11 mos., MZ	0.81 [0.71, 0.91]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta FC3, 8–11 mos., MZ	0.76 [0.65, 0.88]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta F8, 8–11 mos., MZ	0.47 [0.25, 0.69]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta F7, 8–11 mos., MZ	0.41 [0.18, 0.65]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta AF4, 8–11 mos., MZ	0.31 [0.06, 0.57]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Delta AF3, 8–11 mos., MZ	0.36 [0.12, 0.61]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha T6, 8–11 mos., MZ	0.58 [0.39, 0.77]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha T5, 8–11 mos., MZ	0.59 [0.40, 0.77]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha PO4, 8–11 mos., MZ	0.63 [0.46, 0.80]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha PO3, 8–11 mos., MZ	0.74 [0.61, 0.87]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha O2, 8–11 mos., MZ	0.47 [0.25, 0.69]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha O1, 8–11 mos., MZ	0.54 [0.33, 0.74]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha FC4, 8–11 mos., MZ	0.67 [0.51, 0.82]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha FC3, 8–11 mos., MZ	0.69 [0.54, 0.84]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha F8, 8–11 mos., MZ	0.50 [0.28, 0.71]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha F7, 8–11 mos., MZ	0.65 [0.48, 0.81]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha AF4, 8–11 mos., MZ	0.47 [0.25, 0.69]
MCTS, Orekhova 2003, EEG spectral amplitude during visual attention: Alpha AF3, 8–11 mos., MZ	0.47 [0.25, 0.69]
MCTS, Orekhova 2003, EEG mu gravity frequency during visual attention, 8–11 mos., MZ	0.69 [0.53, 0.85]
LTS, Matheny 1980, Task orientation, 9 mos., MZ	0.28 [0.07, 0.49]
LTS, Matheny 1980, Task orientation, 3 mos., MZ	0.55 [0.39, 0.71]
LTS, Davis 2015, Task orientation, 24 mos., MZ	0.30 [0.18, 0.42]
LTS, Davis 2015, Task orientation, 18 mos., MZ	0.37 [0.25, 0.49]
LTS, Davis 2015, Task orientation, 12 mos., MZ	0.40 [0.29, 0.51]
LTS, Davis 2015, Task orientation, 6 mos., MZ	0.45 [0.33, 0.57]
LTS, Matheny 1983, Task orientation, 12–24 mos., MZ	0.49 [0.31, 0.67]
LTS, Matheny 1983, Task orientation, 6–18 mos., MZ	0.53 [0.35, 0.71]
LTS, Matheny 1976, Object orientation, 18–30 mos., MZ	0.90 [0.85, 0.95]
LTS, Matheny 1976, Object orientation, 3–12 mos., MZ	0.82 [0.73, 0.91]
LTS, Matheny 1976, Looking, 18–30 mos., MZ	0.33 [0.07, 0.59]
LTS, Matheny 1976, Looking, 3–12 mos., MZ	0.03 [–0.24, 0.30]
LTS, Matheny 1976, Listening, 18–30 mos., MZ	0.05 [–0.24, 0.34]
LTS, Matheny 1976, Listening, 3–12 mos., MZ	0.65 [0.50, 0.80]
LTS, Matheny 1976, Goal directedness, 18–30 mos., MZ	0.46 [0.23, 0.69]
LTS, Matheny 1976, Goal directedness, 3–12 mos., MZ	0.53 [0.34, 0.72]
LTS, Matheny 1976, Endurance, 18–30 mos., MZ	0.34 [0.08, 0.60]
LTS, Matheny 1976, Endurance, 3–12 mos., MZ	0.62 [0.46, 0.78]
LTS, Matheny 1976, Attention, 18–30 mos., MZ	0.44 [0.21, 0.67]
LTS, Matheny 1976, Attention, 3–12 mos., MZ	0.52 [0.33, 0.71]
JEP, Saudino 2008, Mastery Motivation, 24 mos., MZM	0.84 [0.80, 0.88]
JEP, Saudino 2008, Attention, 24 mos., MZM	0.77 [0.71, 0.83]
JEP, Saudino 2008, Mastery Motivation, 24 mos., MZF	0.78 [0.72, 0.84]
JEP, Saudino 2008, Attention, 24 mos., MZF	0.82 [0.77, 0.87]
CTR_TIP, DiLalla 1996, Task orientation, 9 mos., MZ	0.43 [0.24, 0.62]
CTR_TIP, DiLalla 1996, Task orientation, 7 mos., MZ	0.07 [–0.16, 0.30]
CTR_LTS, Reznick 1997, Visual attentiveness, 24 mos., MZF	0.62 [0.48, 0.76]
CTR_LTS, Reznick 1997, Visual attentiveness, 20 mos., MZF	0.70 [0.58, 0.82]
CTR_LTS, Reznick 1997, Visual attentiveness, 14 mos., MZF	0.55 [0.40, 0.70]
CTR_LTS, Reznick 1997, Visual attentiveness, 24 mos., MZM	0.62 [0.48, 0.76]
CTR_LTS, Reznick 1997, Visual attentiveness, 20 mos., MZM	0.64 [0.50, 0.78]
CTR_LTS, Reznick 1997, Visual attentiveness, 14 mos., MZM	0.52 [0.37, 0.67]
CTR_LTS, Saudino 1996, Task orientation, 24 mos., MZ	0.23 [0.08, 0.38]
CTR_LTS, Saudino 1996, Task orientation, 20 mos., MZ	0.37 [0.24, 0.50]
CTR_LTS, Saudino 1996, Task orientation, 14 mos., MZ	0.17 [0.02, 0.32]
CTR_LTS, Plomin 1993, Task orientation, 14–20 mos., MZ	0.47 [0.30, 0.64]
CTR_LTS, Plomin 1993, Attention/persistence, 14–20 mos., MZ	0.32 [0.13, 0.51]
CTR_LTS, Plomin 1993, Attention/persistence, 14–20 mos., MZ	0.38 [0.20, 0.56]
CTR_LTS, Plomin 1993, Attention/persistence, 20 mos., MZ	0.32 [0.13, 0.51]
CTR_LTS, Plomin 1993, Attention/persistence, 14 mos., MZ	0.38 [0.21, 0.55]
CTR_LTS, Emde 1992, Attention/persistence, 14 mos., MZ	0.38 [0.21, 0.55]
CPP, Goldsmith 1981, Pursuit persistence, 8 mos., MZ	0.42 [0.27, 0.57]
BUTP, Flom 2018, ADHD, 24 mos., MZ	0.74 [0.67, 0.81]
RE Model for MZ Subset	0.60 [0.49, 0.70]

Figure 5A (Continued)

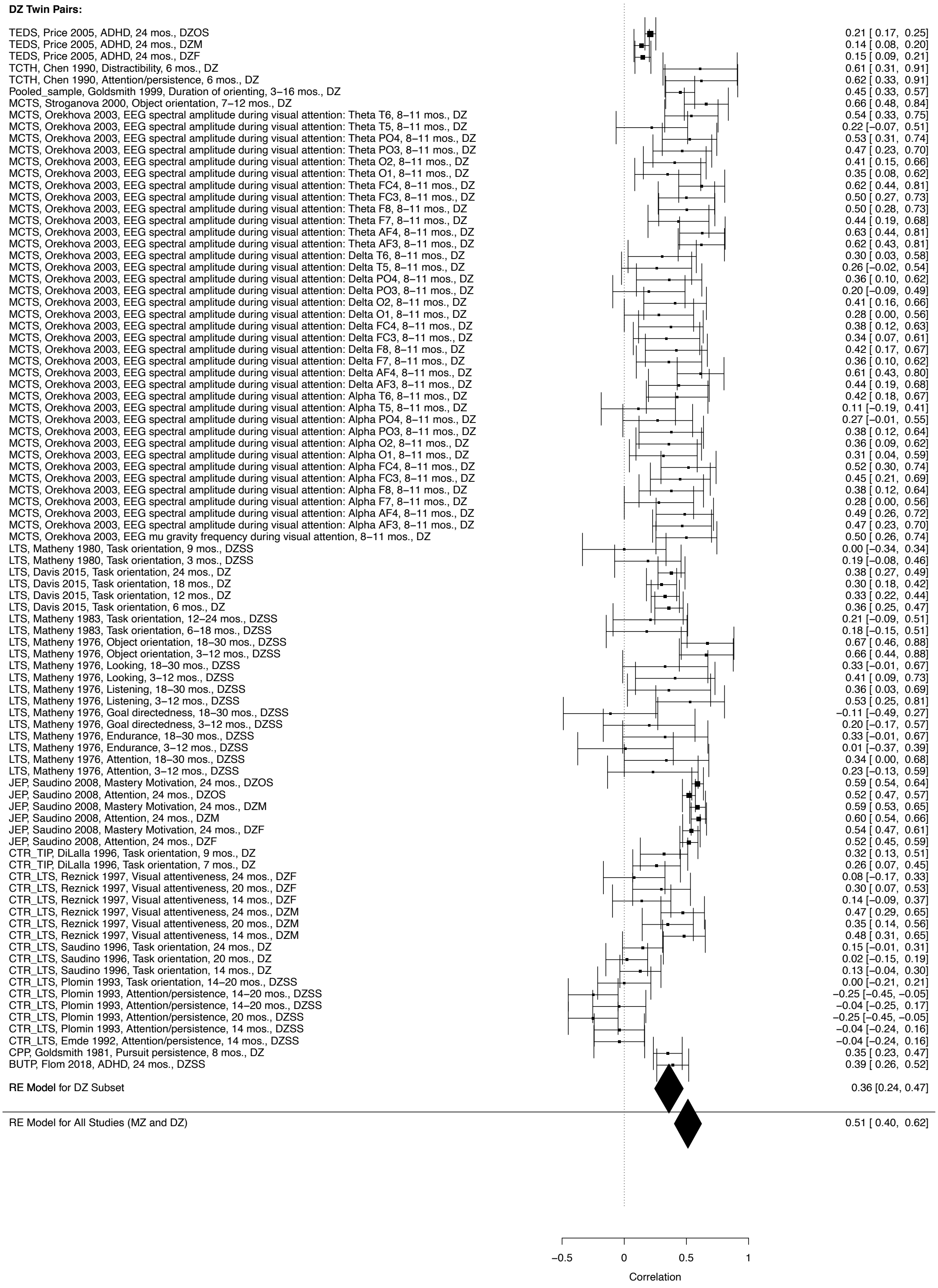


Figure 6A. Psychomotor Functions Forest Plot

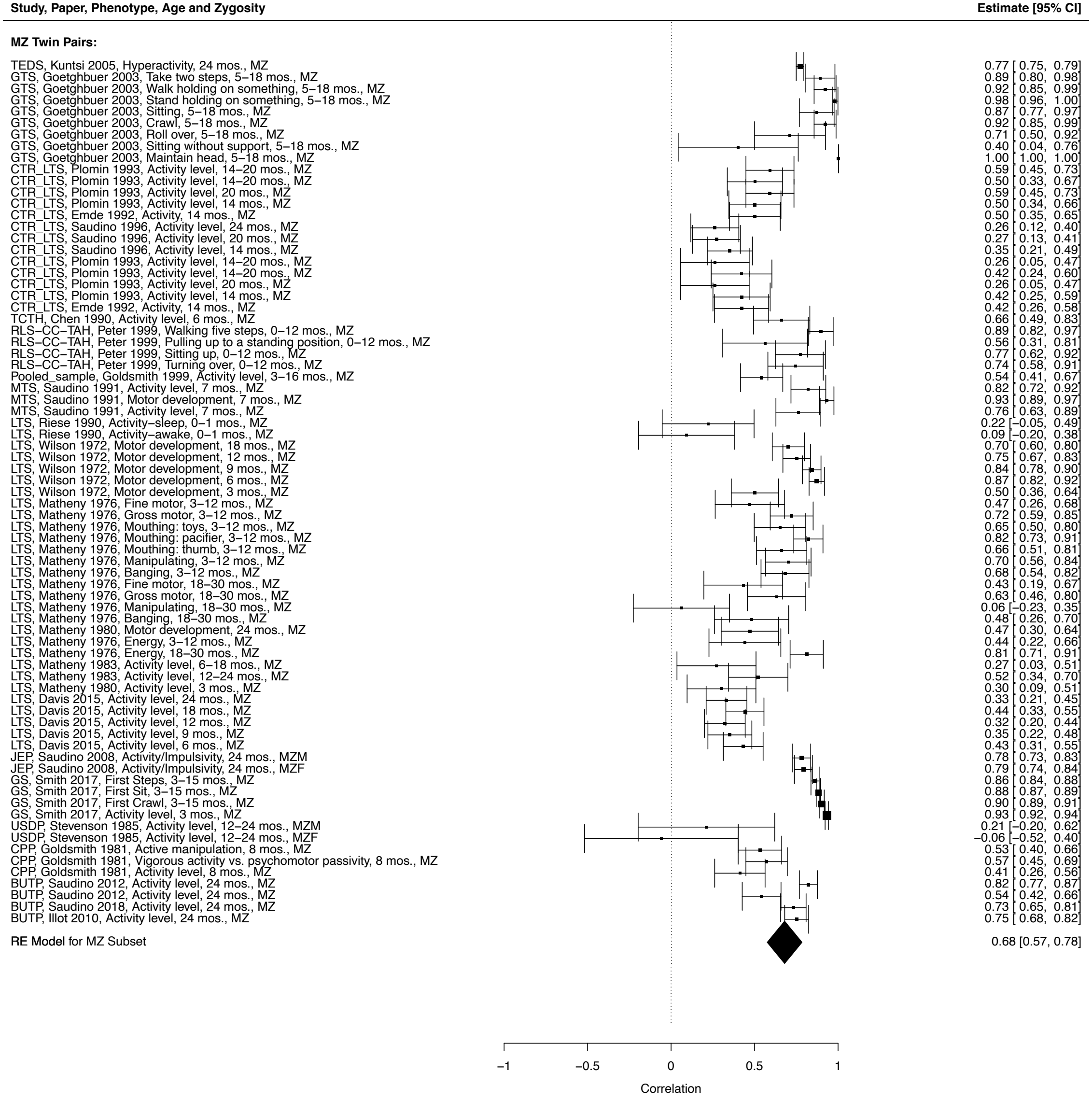


Figure 6A (Continued)

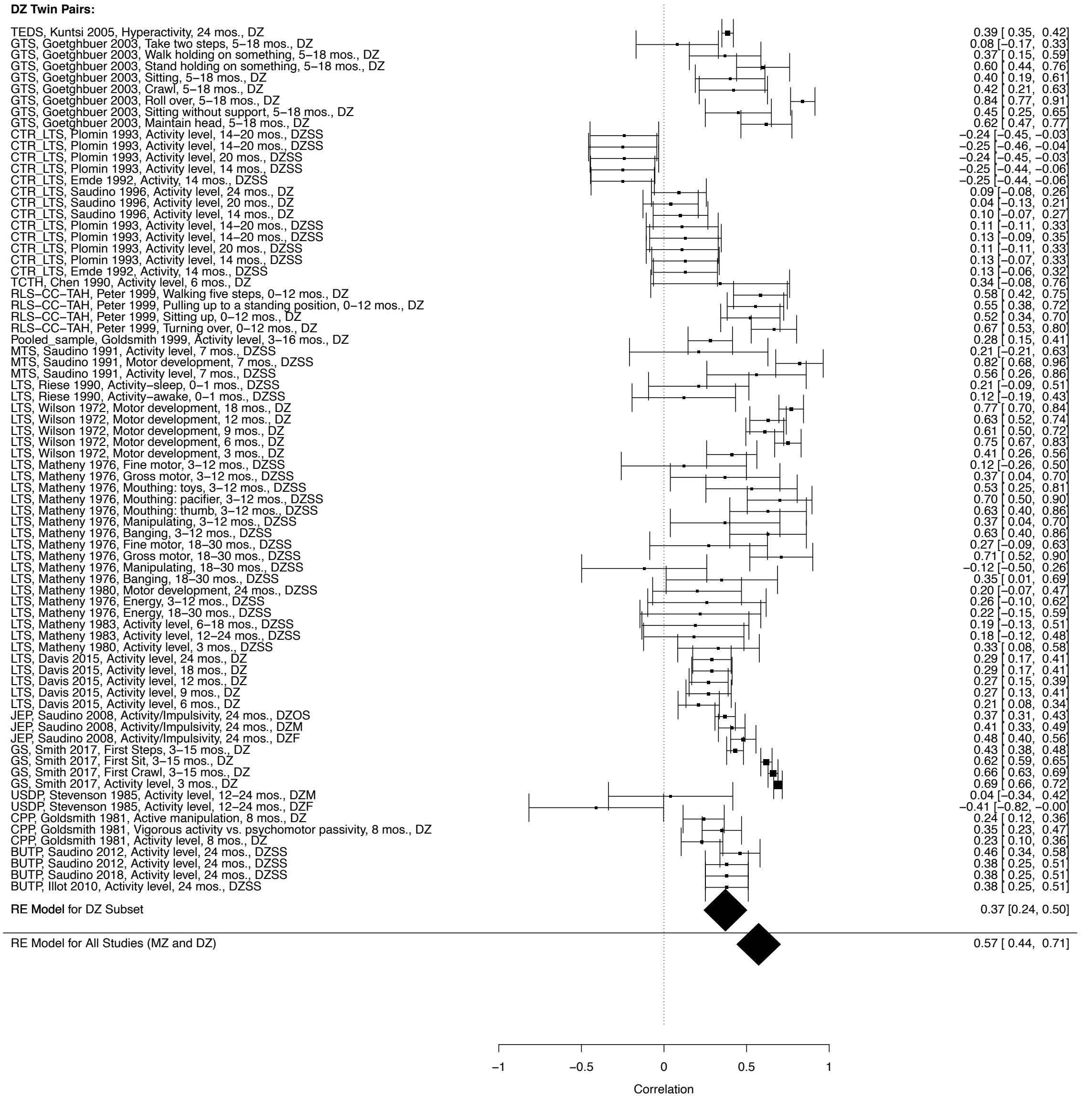
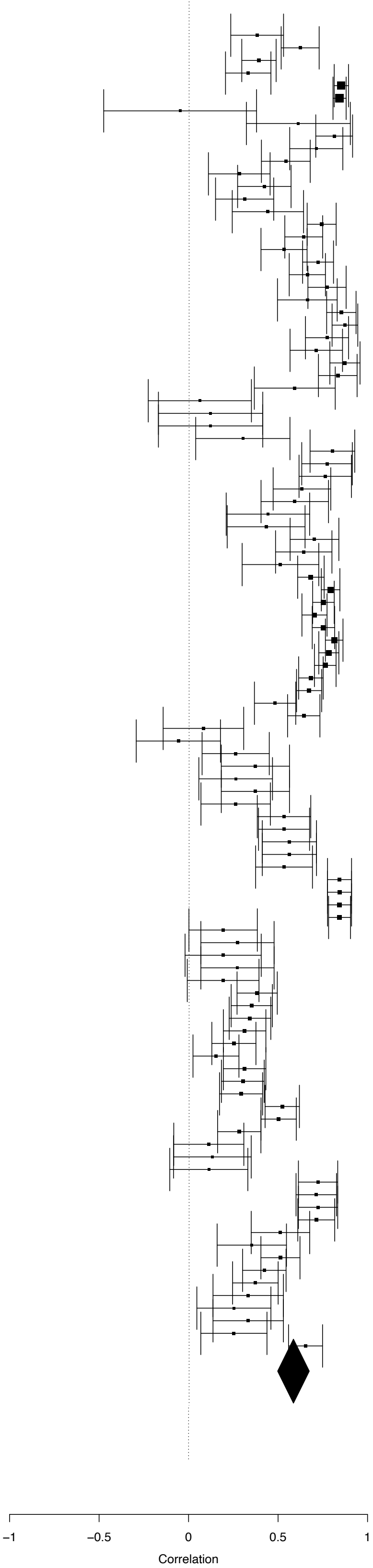


Figure 7A. Emotional Functions Forest Plot



Study, Paper, Phenotype, Age and Zygosity	**Estimate [95% CI]**
MZ Twin Pairs:	
WTP, Gagne 2011, Anger, 12 mos., MZ	0.38 [0.23, 0.53]
WTP, Gagne 2011, Distress to limitations, 12 mos., MZ	0.62 [0.51, 0.73]
WC_BTTYP, Planalp 2017, Positive affect, 12 mos., MZ	0.39 [0.29, 0.49]
WC_BTTYP, Planalp 2017, Positive affect, 6 mos., MZ	0.33 [0.20, 0.46]
WC_BTTYP, Planalp 2017, Smiling and laughter, 6 mos., MZ	0.85 [0.81, 0.89]
WC_BTTYP, Planalp 2017, Smiling and laughter, 12 mos., MZ	0.84 [0.80, 0.88]
USDP, Stevenson 1985, Emotionality, 0–24 mos., MZM	–0.05 [–0.48, 0.38]
USDP, Stevenson 1985, Emotionality, 0–24 mos., MZF	0.61 [0.32, 0.90]
TCTH, Chen 1990, Quality of mood, 6 mos., MZ	0.81 [0.71, 0.91]
TCTH, Chen 1990, Intensity of reaction, 6 mos., MZ	0.71 [0.56, 0.86]
QNTS, Schumann 2017, Negative affect, 18 mos., MZM	0.54 [0.40, 0.68]
QNTS, Schumann 2017, Negative affect, 5 mos., MZM	0.28 [0.11, 0.45]
QNTS, Schumann 2017, Negative affect, 5 mos., MZF	0.42 [0.27, 0.57]
QNTS, Schumann 2017, Negative affect, 18 mos., MZF	0.31 [0.15, 0.47]
QNTS, Ouellet–Morin 2008, Cortisol reactivity, 19 mos., MZ	0.44 [0.24, 0.64]
Pooled_sample, Goldsmith 1999, Positive affect, 3–16 mos., MZ	0.74 [0.66, 0.82]
Pooled_sample, Goldsmith 1999, Negative affect, 3–16 mos., MZ	0.64 [0.53, 0.75]
Pooled_sample, Goldsmith 1999, Resistance to soothing, 3–16 mos., MZ	0.53 [0.40, 0.66]
Pooled_sample, Goldsmith 1999, Smiling and laughter, 3–16 mos., MZ	0.72 [0.63, 0.81]
Pooled_sample, Goldsmith 1999, Distress to limitations, 3–16 mos., MZ	0.66 [0.56, 0.76]
NTR_MBF–QCCH, Bokhorst 2003, Temperamental reactivity, 12–14 mos., MZ	0.77 [0.66, 0.88]
MCTS, Stroganova 2000, Nonoriented discharges, 7–12 mos., MZ	0.66 [0.49, 0.83]
MCTS, Stroganova 2000, Moderate tension, 7–12 mos., MZ	0.85 [0.77, 0.93]
MCTS, Stroganova 2000, Low tension, 7–12 mos., MZ	0.87 [0.80, 0.94]
MCTS, Stroganova 2000, High tension, 7–12 mos., MZ	0.77 [0.65, 0.89]
MCTS, Stroganova 2000, Fear, 7–12 mos., MZ	0.71 [0.56, 0.86]
LTS, Matheny 1989, Emotional tone, 24 mos., MZ	0.87 [0.79, 0.95]
LTS, Matheny 1989, Emotional tone, 18 mos., MZ	0.83 [0.72, 0.94]
LTS, Matheny 1989, Emotional tone, 12 mos., MZ	0.59 [0.36, 0.82]
LTS, Riese 1990, Resistance to soothing, 0–1 mos., MZ	0.06 [–0.23, 0.35]
LTS, Riese 1990, Reactivity, 0–1 mos., MZ	0.12 [–0.17, 0.41]
LTS, Riese 1990, Reactivity, 0–1 mos., MZ	0.12 [–0.17, 0.41]
LTS, Riese 1990, Irritability, 0–1 mos., MZ	0.30 [0.04, 0.56]
LTS, Matheny 1989, Fearfulness, 24 mos., MZ	0.80 [0.68, 0.92]
LTS, Matheny 1989, Fearfulness, 18 mos., MZ	0.77 [0.63, 0.91]
LTS, Matheny 1989, Fearfulness, 12 mos., MZ	0.76 [0.61, 0.91]
LTS, Matheny 1976, Tension, 3–12 mos., MZ	0.63 [0.47, 0.79]
LTS, Matheny 1976, Tension, 18–30 mos., MZ	0.59 [0.40, 0.78]
LTS, Matheny 1976, Reactivity, 18–30 mos., MZ	0.44 [0.21, 0.67]
LTS, Matheny 1976, Reactivity, 3–12 mos., MZ	0.43 [0.21, 0.65]
LTS, Matheny 1976, Fearfulness, 3–12 mos., MZ	0.70 [0.56, 0.84]
LTS, Matheny 1976, Emotional tone, 3–12 mos., MZ	0.64 [0.48, 0.80]
LTS, Matheny 1976, Emotional tone, 18–30 mos., MZ	0.51 [0.30, 0.72]
JEP, Saudino 2008, Negative affect, 24 mos., MZM	0.68 [0.61, 0.75]
JEP, Saudino 2008, Internalizing, 24 mos., MZM	0.79 [0.74, 0.84]
JEP, Saudino 2008, General anxiety, 24 mos., MZM	0.75 [0.69, 0.81]
JEP, Saudino 2008, Depression withdrawal, 24 mos., MZM	0.70 [0.63, 0.77]
JEP, Saudino 2008, Negative affect, 24 mos., MZF	0.75 [0.69, 0.81]
JEP, Saudino 2008, Internalizing, 24 mos., MZF	0.81 [0.76, 0.86]
JEP, Saudino 2008, General anxiety, 24 mos., MZF	0.78 [0.72, 0.84]
JEP, Saudino 2008, Depression withdrawal, 24 mos., MZF	0.76 [0.70, 0.82]
ECLS–B, Jackson 2016, Moody/unusual, 24 mos., MZ	0.68 [0.61, 0.75]
ECLS–B, Jackson 2016, Demanding/angry, 24 mos., MZ	0.67 [0.60, 0.74]
ECLS–B, Roisman 2006, Fussiness and demanding behaviour, 9 mos., MZ	0.48 [0.36, 0.60]
ECLS–B, Roisman 2006, Positive and negative affect, 9 mos., MZ	0.64 [0.55, 0.73]
CTR_TIP, DiLalla 1996, Negative affect, 7 mos., MZ	0.08 [–0.14, 0.30]
CTR_TIP, DiLalla 1996, Negative affect, 9 mos., MZ	–0.06 [–0.30, 0.18]
CTR_LTS, Emde 1992, Frustration, 14 mos., MZ	0.26 [0.07, 0.45]
CTR_LTS, Plomin 1993, Reactivity, 14–20 mos., MZ	0.37 [0.18, 0.56]
CTR_LTS, Plomin 1993, Reactivity, 14–20 mos., MZ	0.26 [0.05, 0.47]
CTR_LTS, Plomin 1993, Reactivity, 20 mos., MZ	0.37 [0.18, 0.56]
CTR_LTS, Plomin 1993, Reactivity, 14 mos., MZ	0.26 [0.07, 0.45]
CTR_LTS, Plomin 1993, Positive hedonic tone, 14 mos., MZ	0.53 [0.38, 0.68]
CTR_LTS, Emde 1992, Positive hedonic tone, 14 mos., MZ	0.53 [0.39, 0.67]
CTR_LTS, Plomin 1993, Positive hedonic tone, 20 mos., MZ	0.56 [0.41, 0.71]
CTR_LTS, Plomin 1993, Positive affect, 14–20 mos., MZ	0.56 [0.41, 0.71]
CTR_LTS, Plomin 1993, Positive affect, 14–20 mos., MZ	0.53 [0.37, 0.69]
CTR_LTS, Plomin 1993, Positive affect, 14–20 mos., MZ	0.84 [0.77, 0.91]
CTR_LTS, Plomin 1993, Positive affect, 14–20 mos., MZ	0.84 [0.77, 0.91]
CTR_LTS, Plomin 1993, Positive affect, 20 mos., MZ	0.84 [0.77, 0.91]
CTR_LTS, Plomin 1993, Positive affect, 14 mos., MZ	0.84 [0.78, 0.90]
CTR_LTS, Emde 1992, Overall mood, 14 mos., MZ	0.19 [–0.00, 0.38]
CTR_LTS, Plomin 1993, Affect, 14–20 mos., MZ	0.27 [0.07, 0.47]
CTR_LTS, Plomin 1993, Affect, 14–20 mos., MZ	0.19 [–0.02, 0.40]
CTR_LTS, Plomin 1993, Overall mood, 20 mos., MZ	0.27 [0.07, 0.47]
CTR_LTS, Plomin 1993, Affect, 14 mos., MZ	0.19 [–0.01, 0.39]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.38 [0.27, 0.49]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.35 [0.23, 0.47]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.34 [0.22, 0.46]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.31 [0.19, 0.43]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.25 [0.13, 0.37]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.15 [0.02, 0.28]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.31 [0.19, 0.43]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.30 [0.18, 0.42]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.29 [0.17, 0.41]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.52 [0.42, 0.62]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.50 [0.40, 0.60]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.28 [0.16, 0.40]
CTR_LTS, Emde 1992, Negative hedonic tone, 14 mos., MZ	0.11 [–0.09, 0.31]
CTR_LTS, Plomin 1993, Negative affect, 14–20 mos., MZ	0.13 [–0.09, 0.35]
CTR_LTS, Plomin 1993, Negative affect, 14–20 mos., MZ	0.11 [–0.11, 0.33]
CTR_LTS, Plomin 1993, Negative affect, 14–20 mos., MZ	0.72 [0.61, 0.83]
CTR_LTS, Plomin 1993, Negative affect, 14–20 mos., MZ	0.71 [0.60, 0.82]
CTR_LTS, Plomin 1993, Negative affect, 20 mos., MZ	0.72 [0.61, 0.83]
CTR_LTS, Plomin 1993, Negative affect, 14 mos., MZ	0.71 [0.61, 0.81]
CTR_LTS, Plomin 1993, Emotionality, 14–20 mos., MZ	0.51 [0.35, 0.67]
CTR_LTS, Plomin 1993, Emotionality, 14–20 mos., MZ	0.35 [0.16, 0.54]
CTR_LTS, Schmitz 1999, Emotionality, 20 mos., MZ	0.51 [0.40, 0.62]
CTR_LTS, Schmitz 1999, Emotionality, 14 mos., MZ	0.42 [0.30, 0.54]
CTR_LTS, Schmitz 1999, Emotionality, 24 mos., MZ	0.37 [0.24, 0.50]
CTR_LTS, Plomin 1993, Affect, 14–20 mos., MZ	0.33 [0.13, 0.53]
CTR_LTS, Plomin 1993, Affect, 14–20 mos., MZ	0.25 [0.04, 0.46]
CTR_LTS, Plomin 1993, Affect, 20 mos., MZ	0.33 [0.13, 0.53]
CTR_LTS, Emde 1992, Affect, 14 mos., MZ	0.25 [0.07, 0.43]
BUTP, Micalizzi 2016, Affective problems, 24 mos., MZ	0.65 [0.56, 0.74]
RE Model for MZ Subset	0.58 [0.49, 0.67]

Figure 7A (Continued)

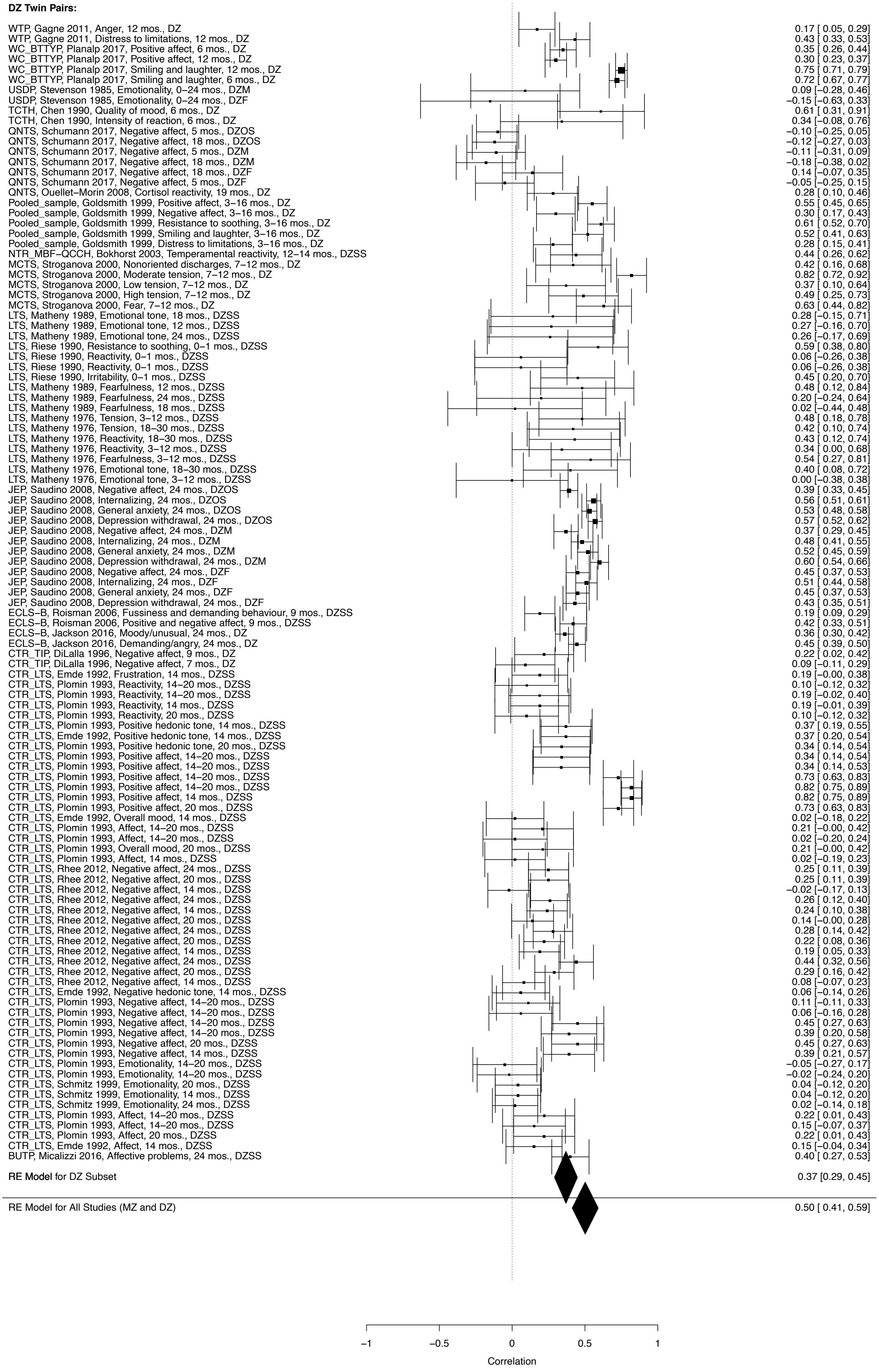


Figure 8A. Basic Cognitive Functions Forest Plot

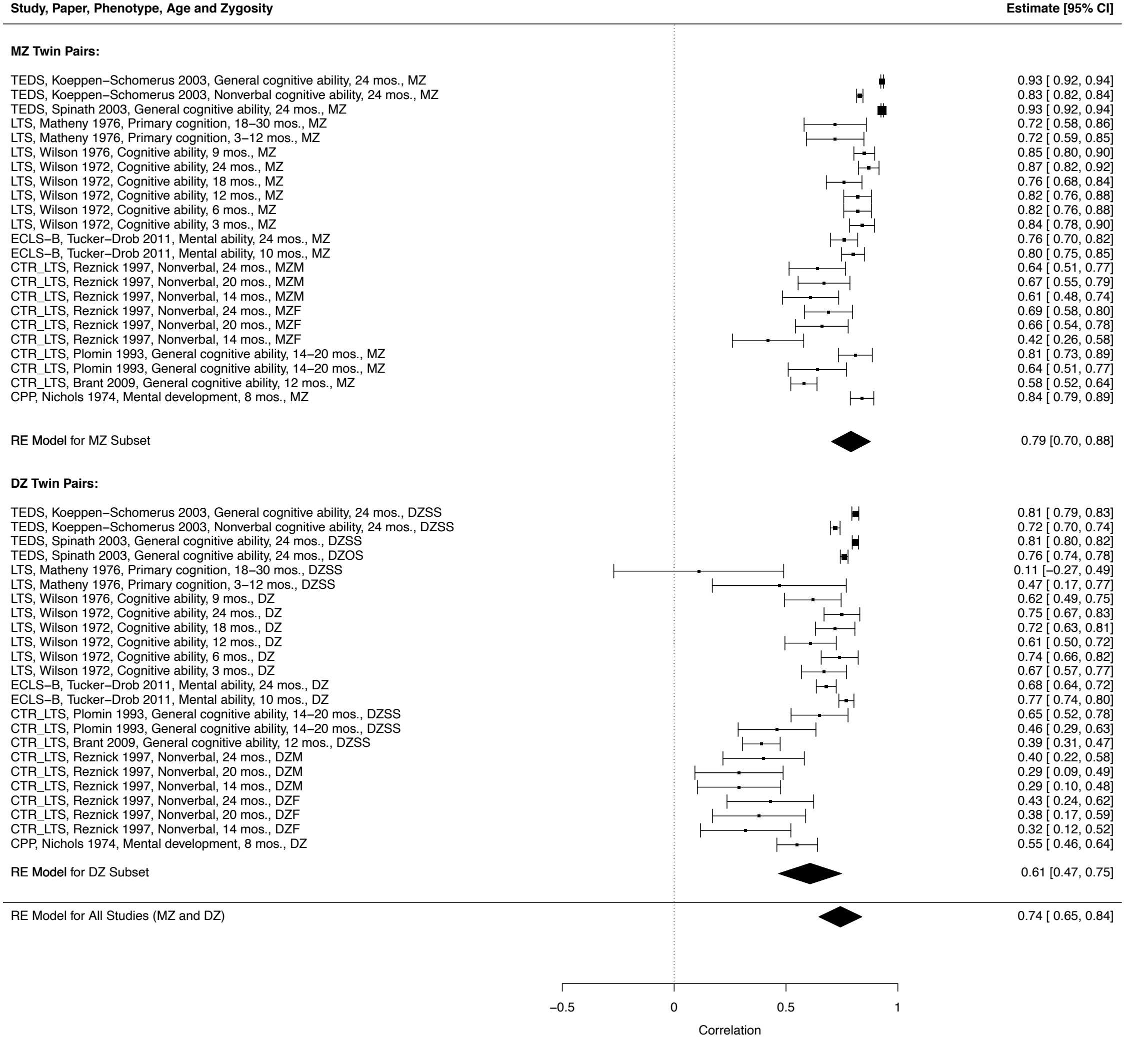


Figure 9A. Mental Functions of Language Forest Plot

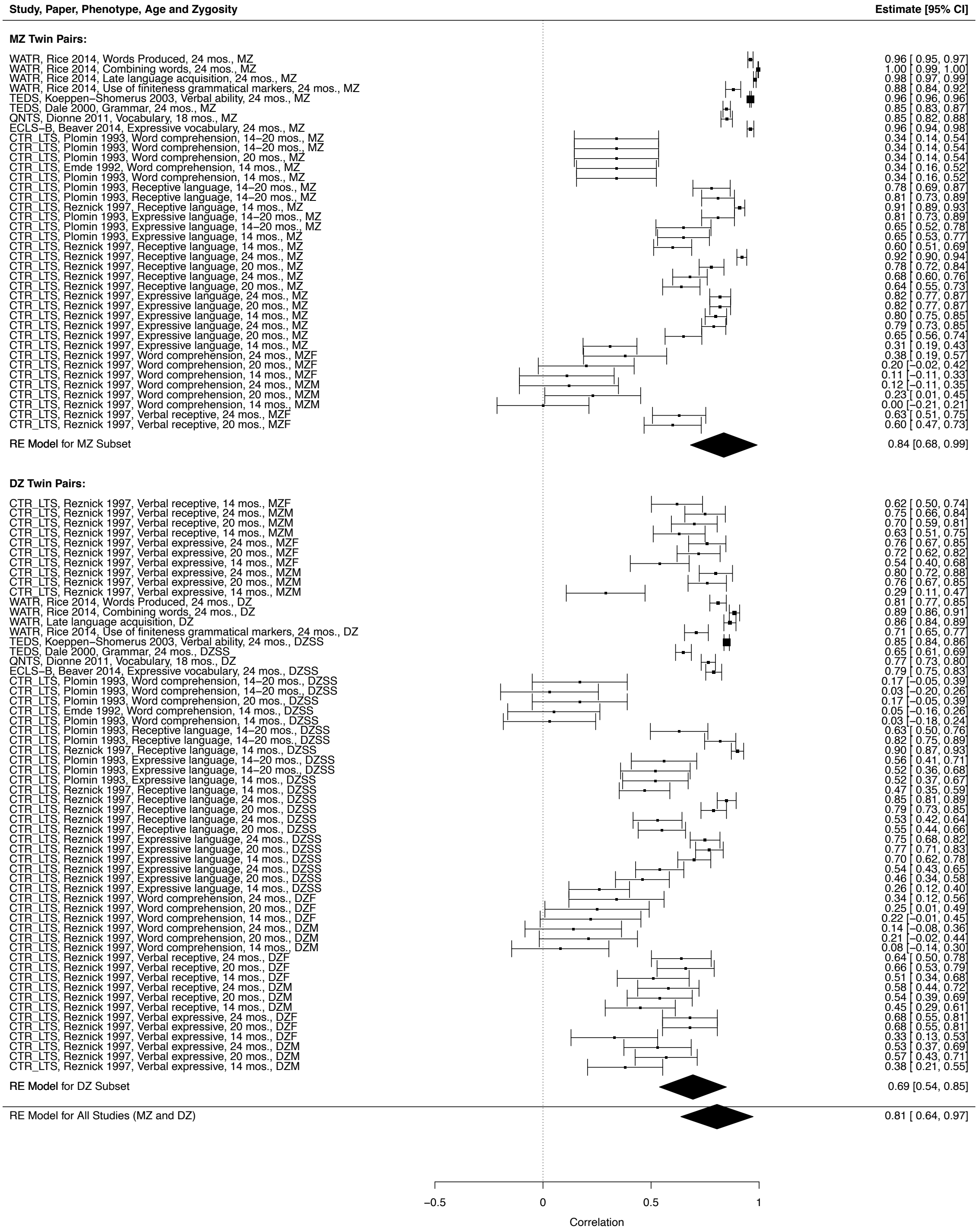


Figure 10A. Growth Maintenance Functions Forest Plot

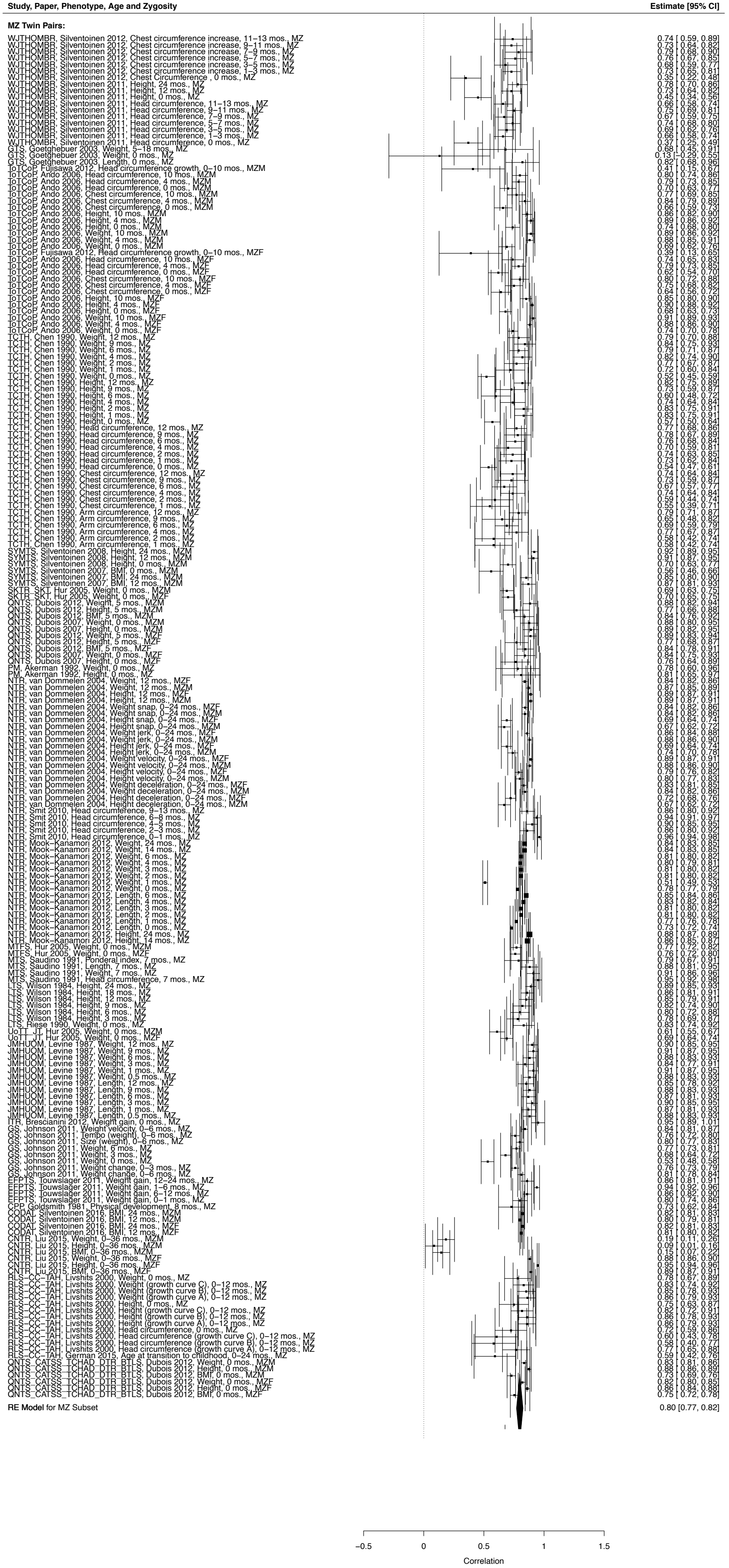


Figure 10A (Continued)

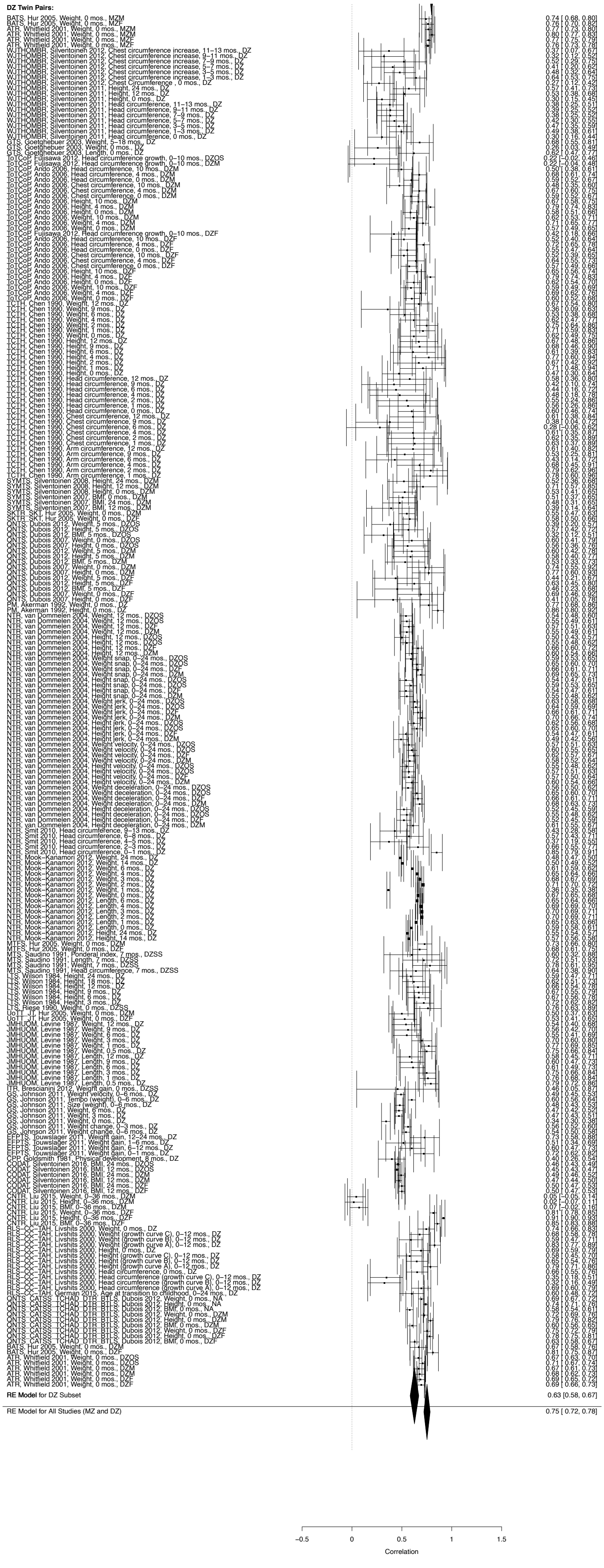


Figure 11A. Basic Interpersonal Interactions Forest Plot

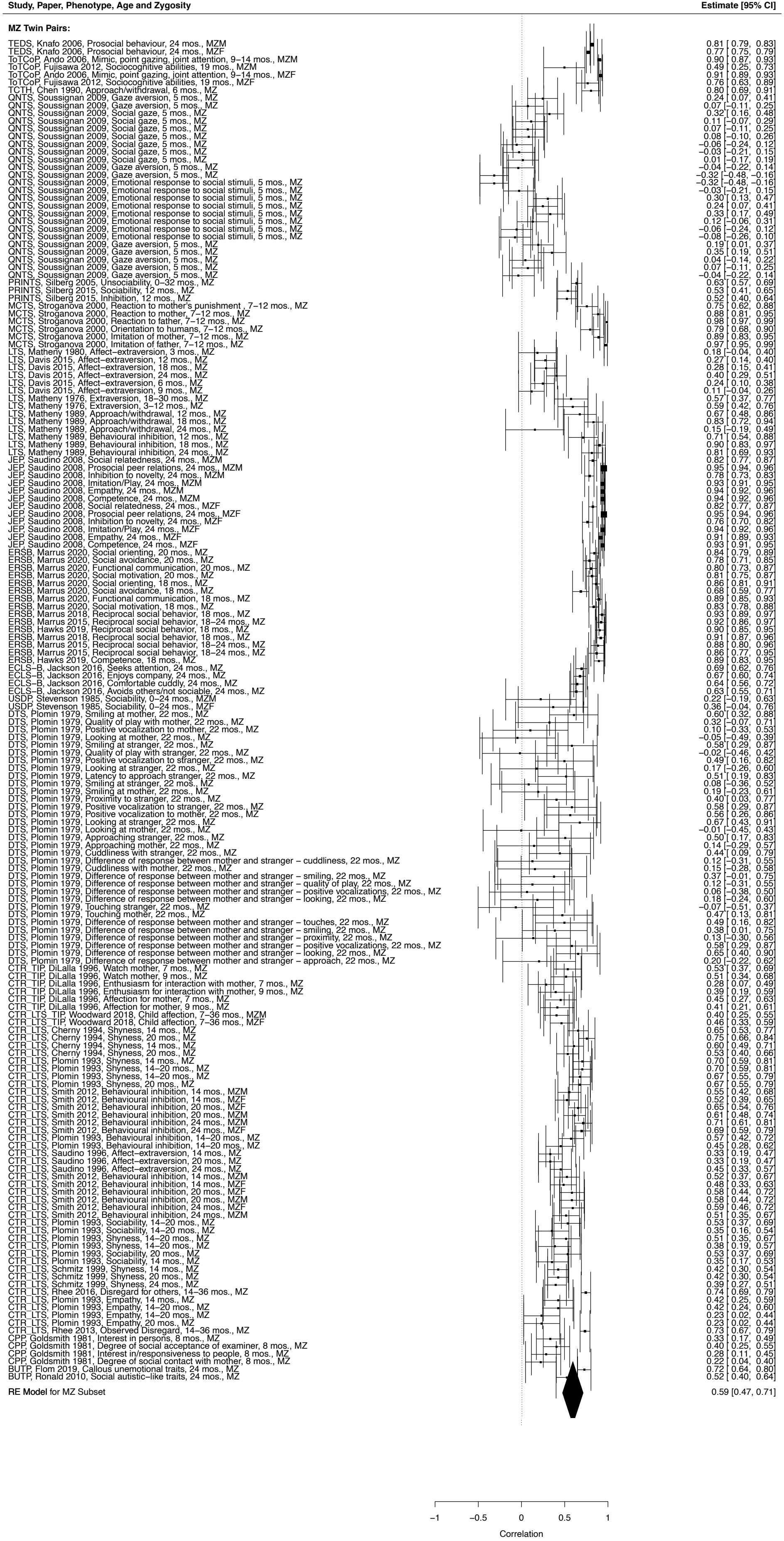


Figure 11A (Continued)

Figure 12A. Complex Interpersonal Interactions Forest Plot

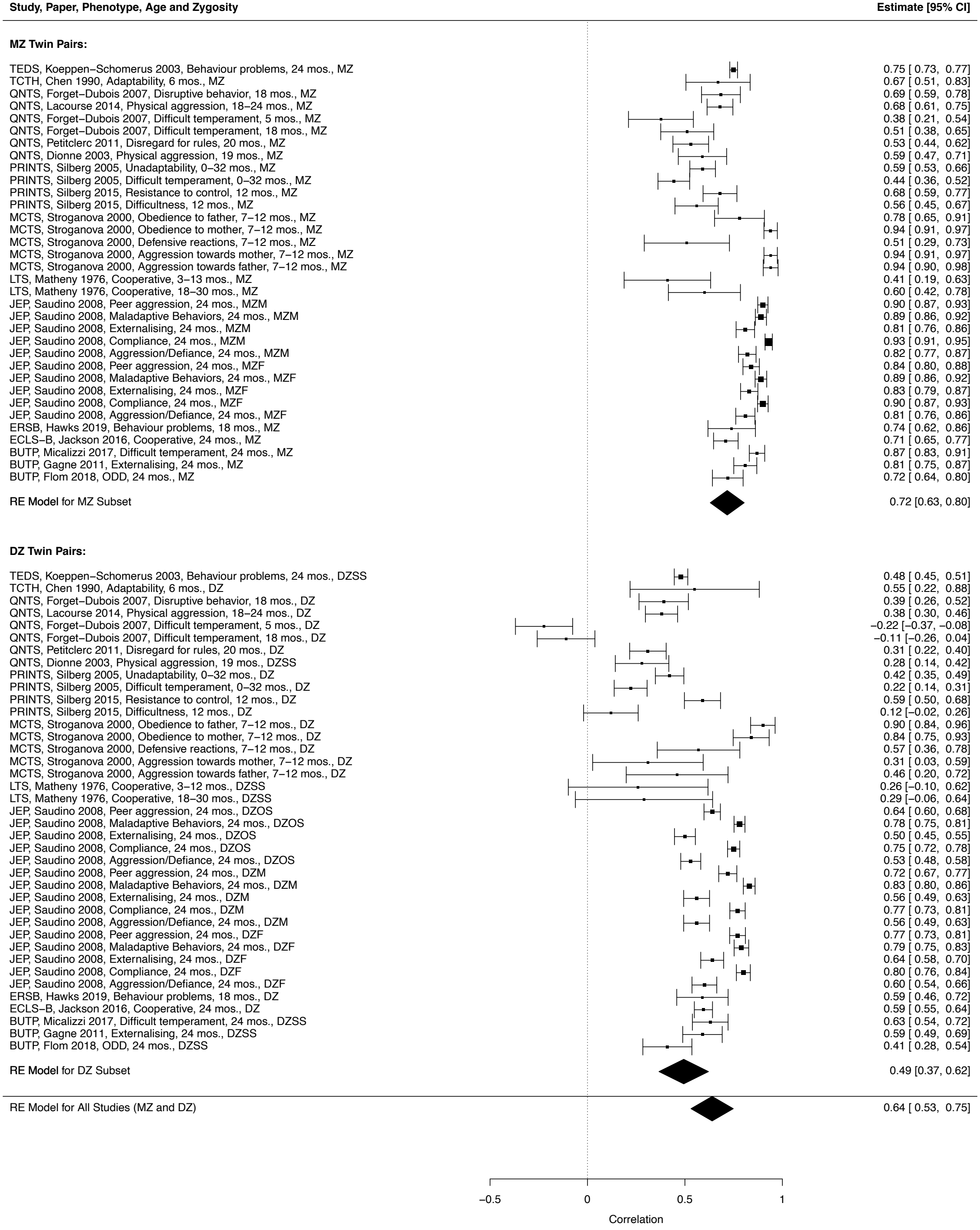


Figure 13A. Family Relationships Forest Plot

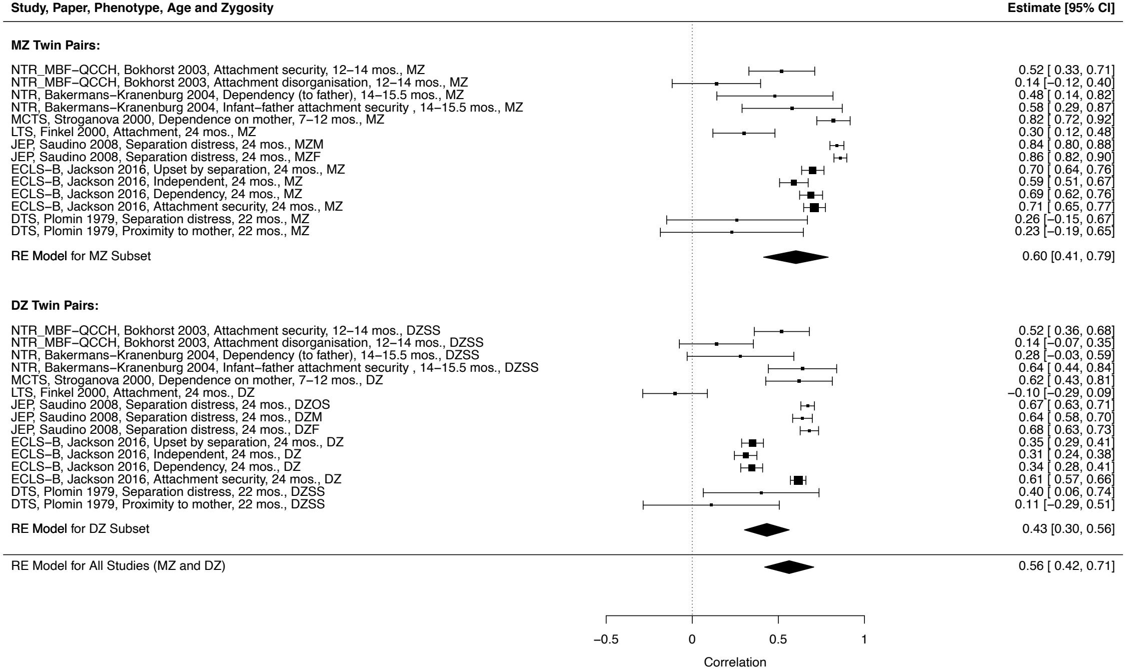


Figure 14A. Funnel Plots of Association Between Monozygotic Twin Correlation (r_{MZ}) and Standard Error in Phenotype Categories With ≥ 10 Estimates

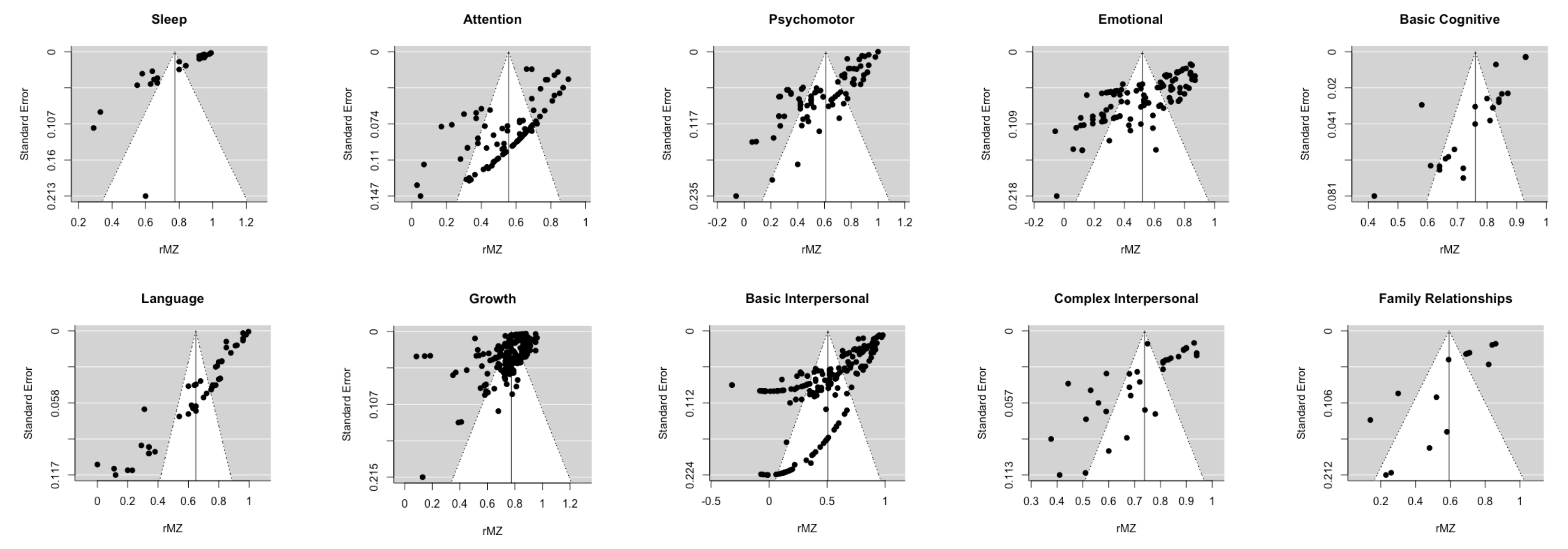
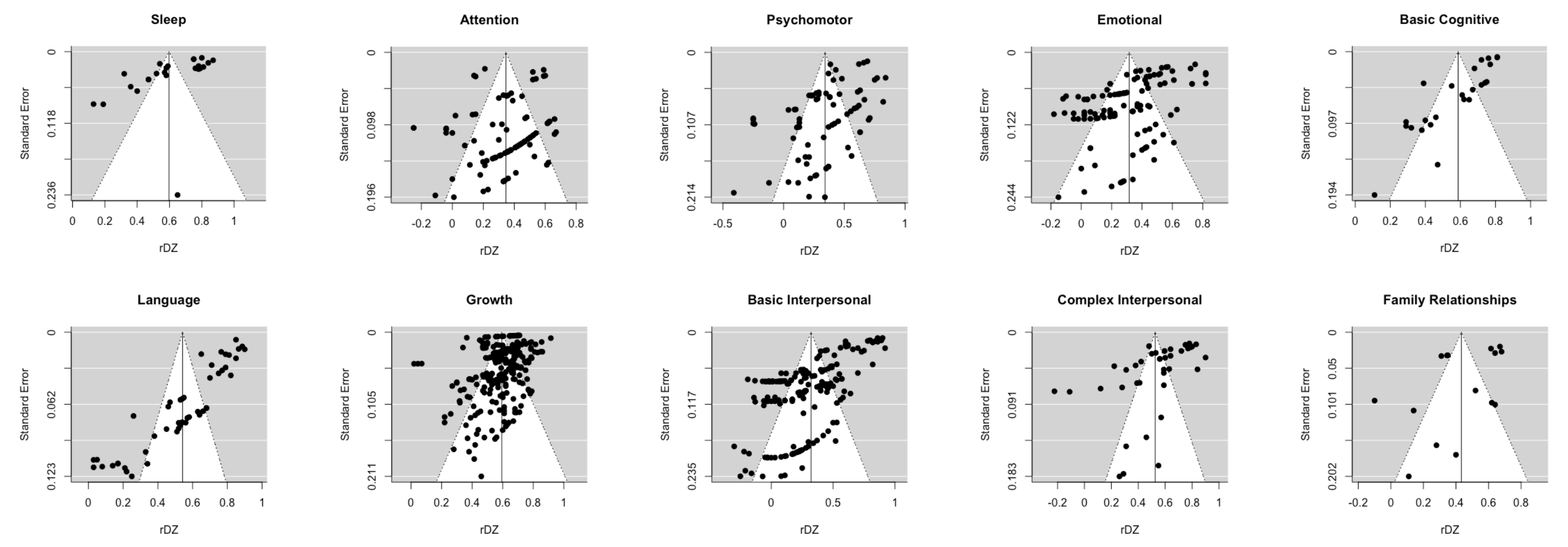


Figure 15. Funnel Plots of Association Between Dizygotic Twin Correlation (r_{DZ}) and Standard Error in Phenotype Categories With ≥ 10 Estimates



Note. r_{DZ} = dizygotic twin correlation

Figure 16A. Funnel Plots of Association Between Heritability (h^2) and Standard Error in Phenotype Categories With ≥ 10 Estimates

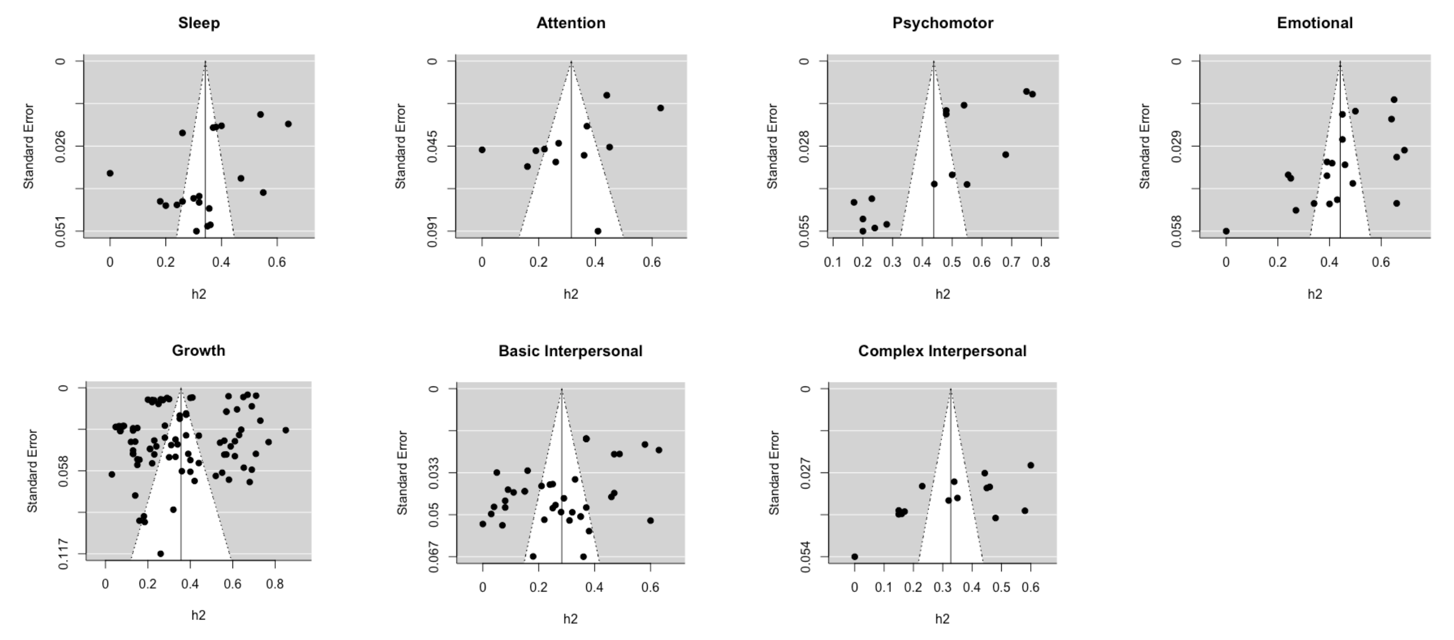


Figure 17A. Funnel Plots of Association Between Shared Environment (c^2) and Standard Error in Phenotype Categories With ≥ 10 Estimates

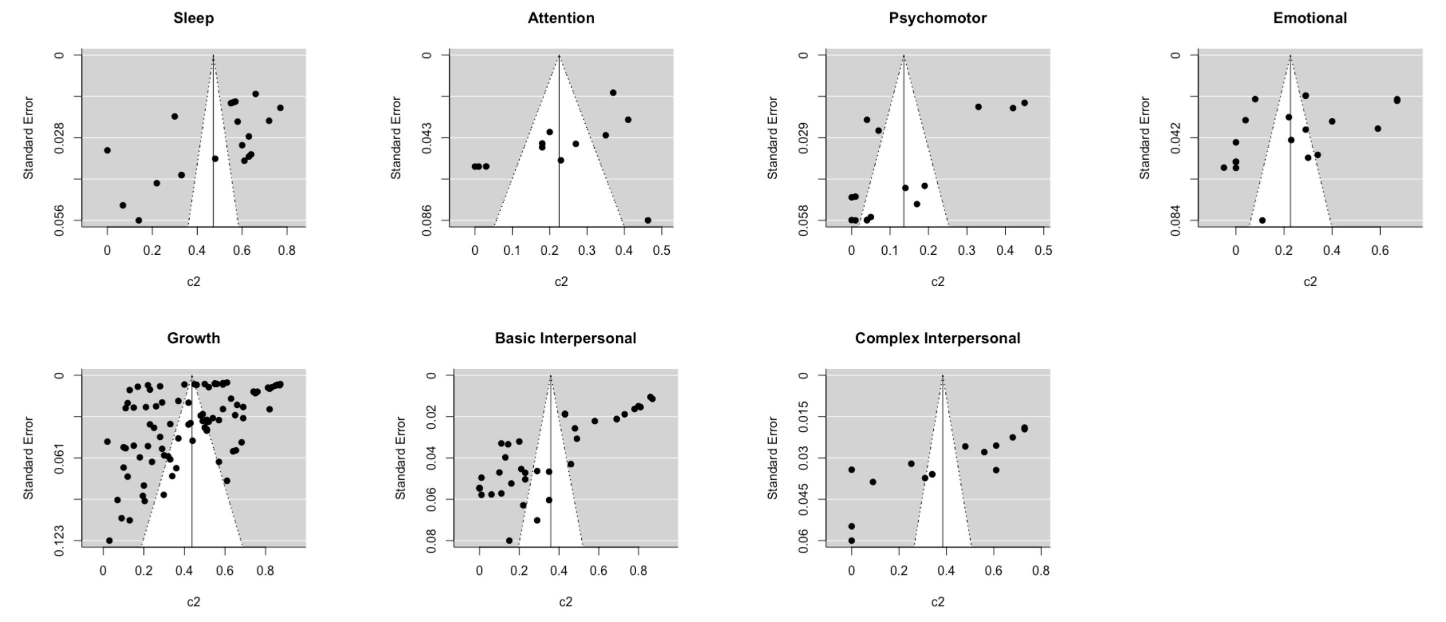


Figure 18A. Funnel Plots of Association Between Nonshared Environment (e^2) and Standard Error in Phenotype Categories With ≥ 10 Estimates

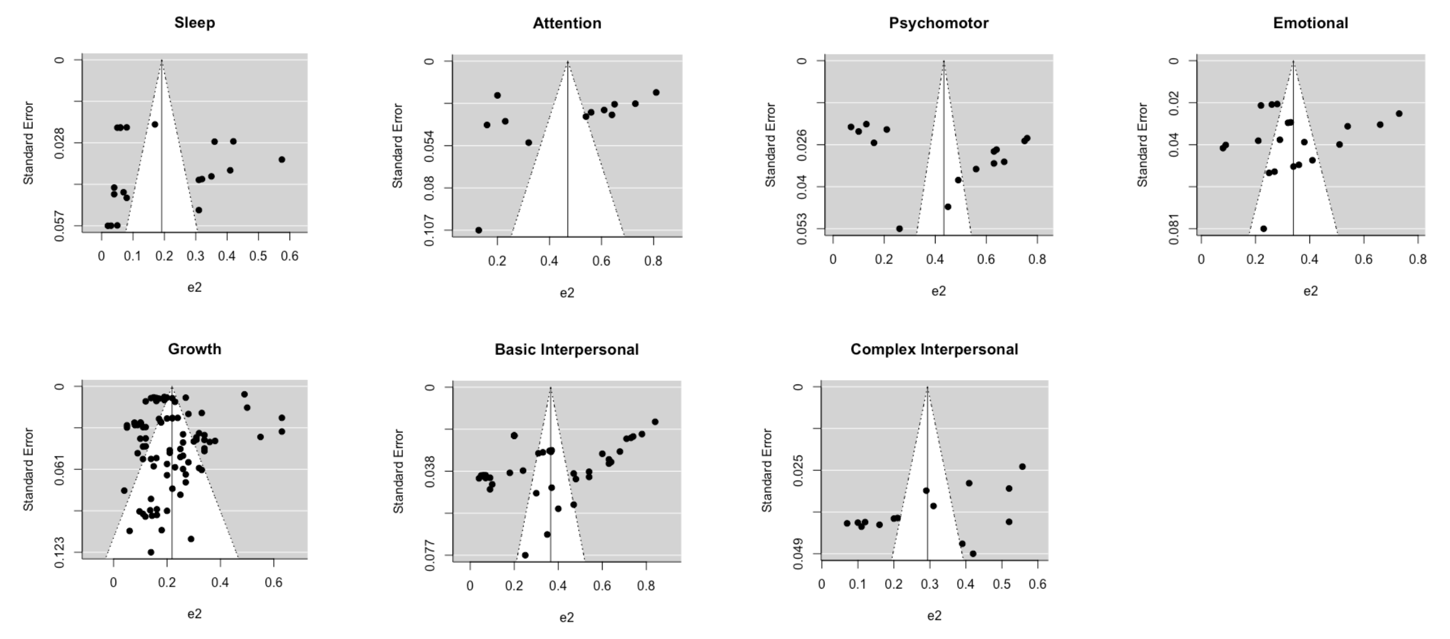


Figure 19A. Bar Plot of Quality Assessment Scores

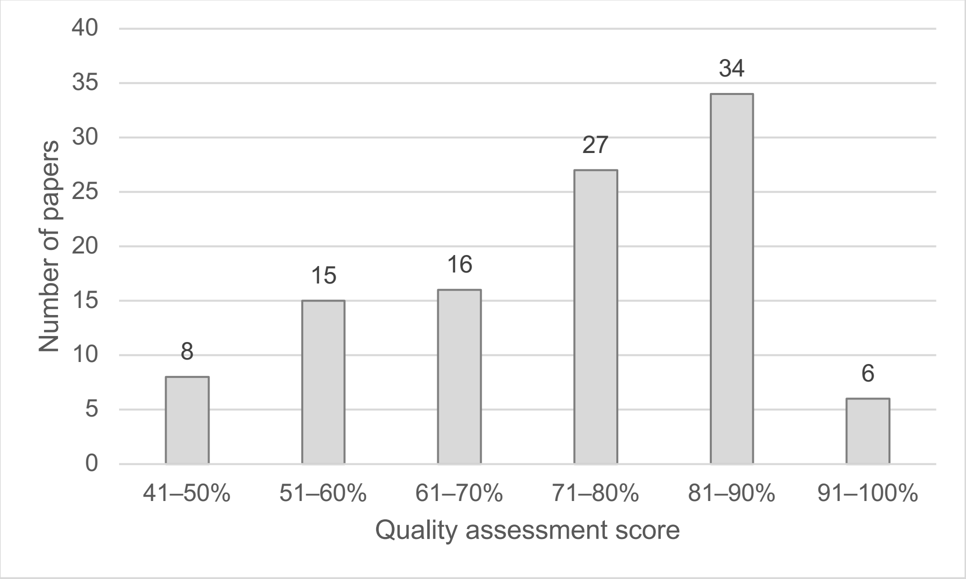


Figure 20A. Sustaining Attention Forest Plot

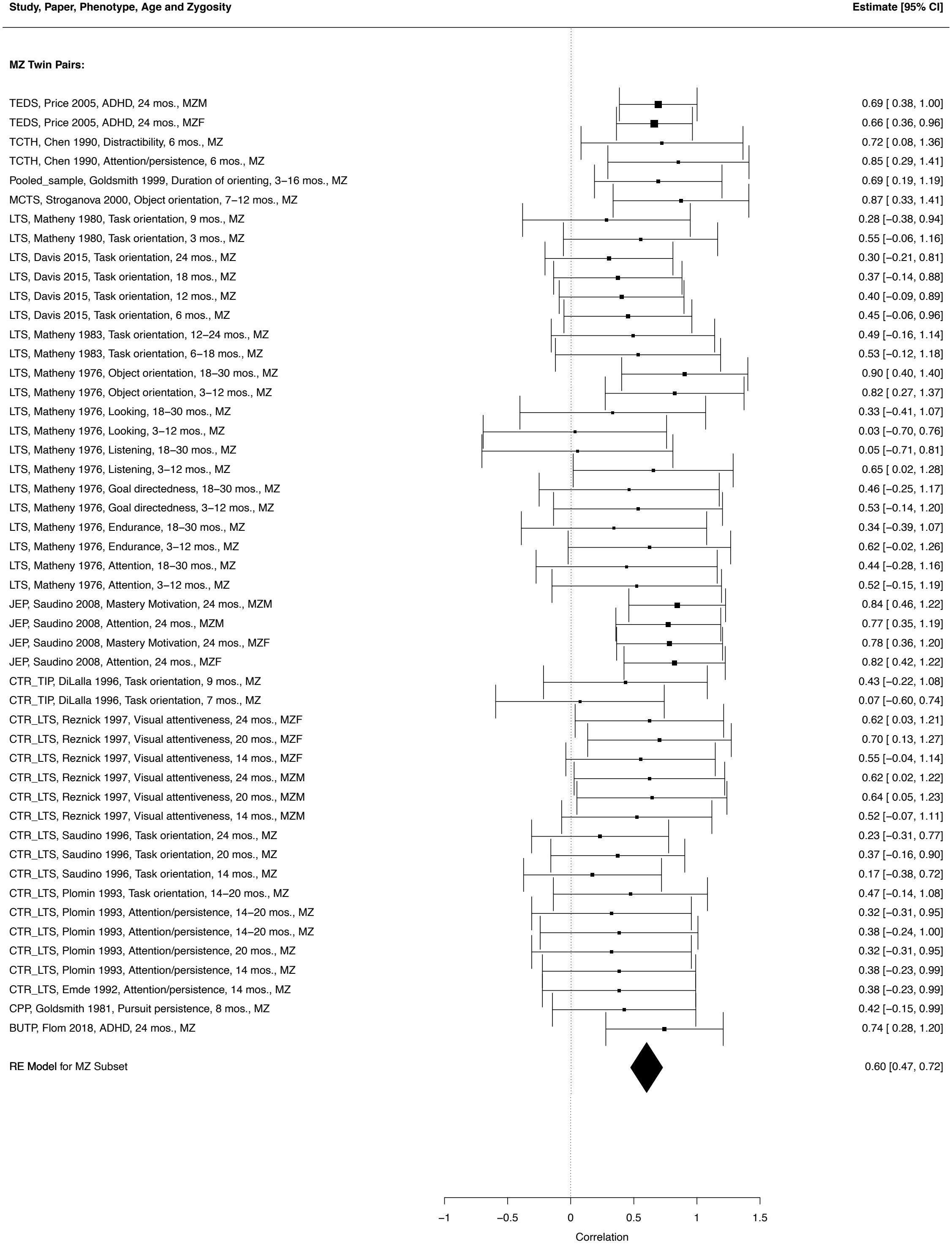


Figure 20A (Continued)

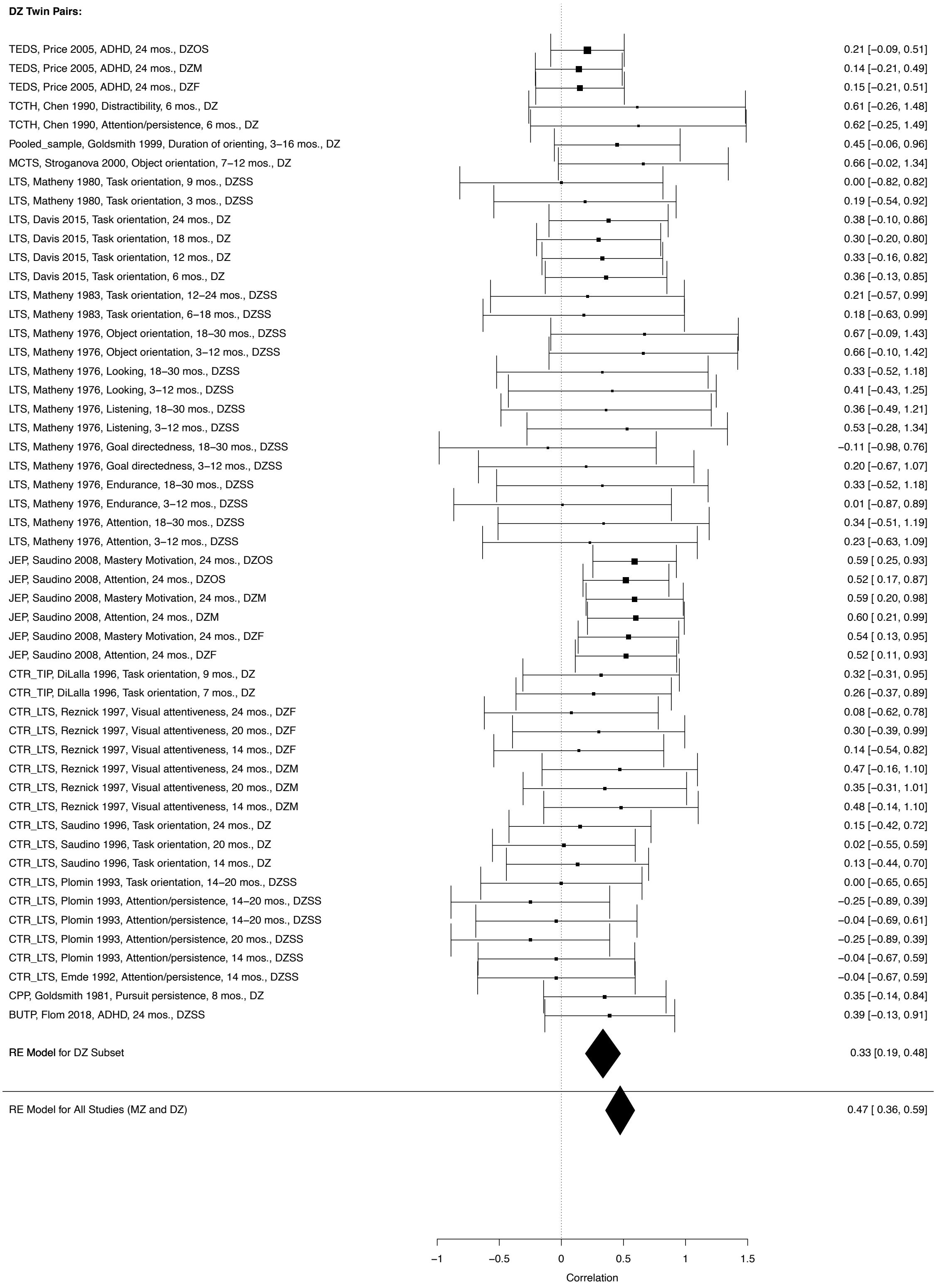
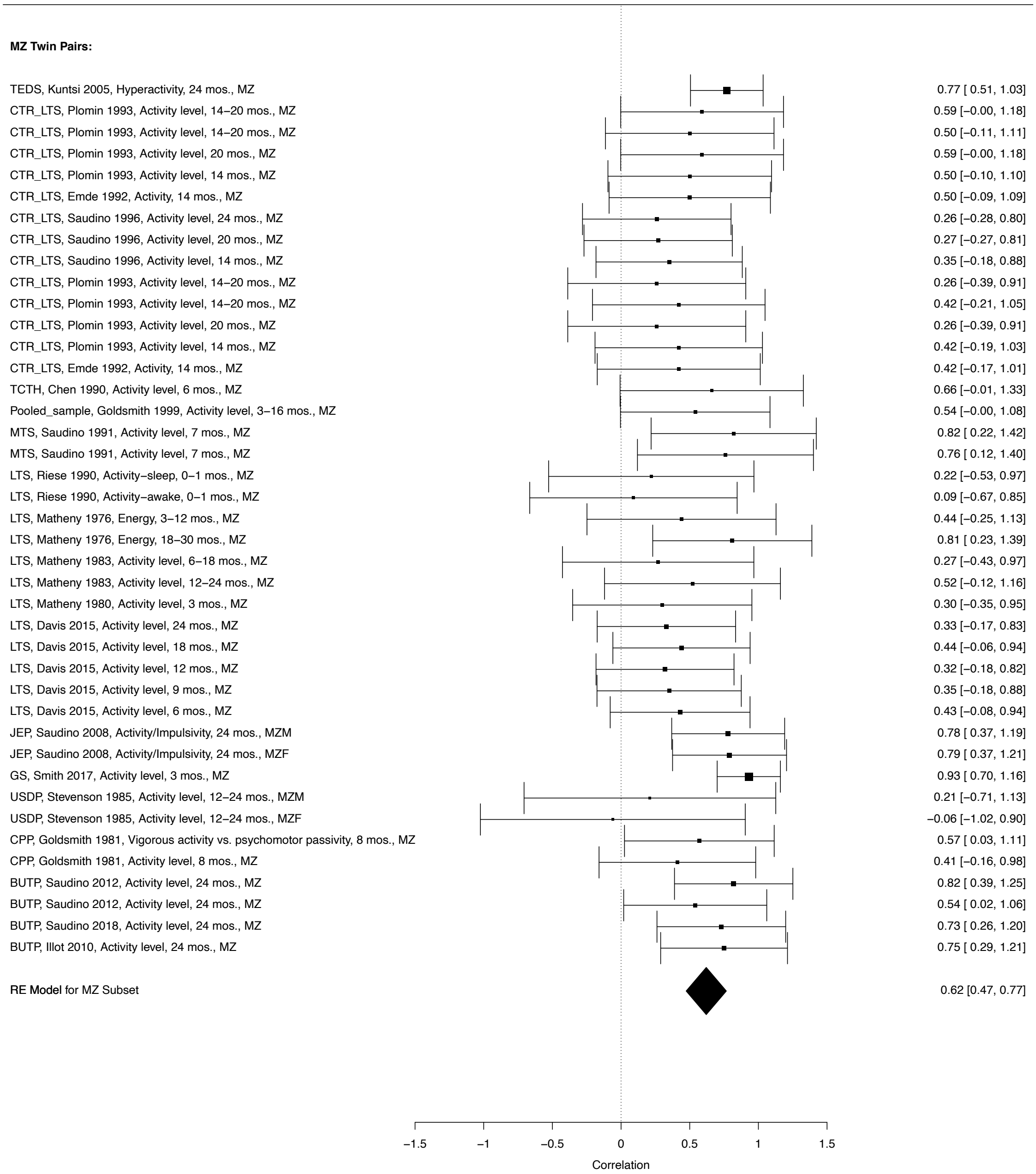


Figure 21A. Psychomotor Control Forest Plot



Study, Paper, Phenotype, Age and Zygosity	Estimate [95% CI]
MZ Twin Pairs:	
TEDS, Kuntsi 2005, Hyperactivity, 24 mos., MZ	0.77 [0.51, 1.03]
CTR_LTS, Plomin 1993, Activity level, 14−20 mos., MZ	0.59 [−0.00, 1.18]
CTR_LTS, Plomin 1993, Activity level, 14−20 mos., MZ	0.50 [−0.11, 1.11]
CTR_LTS, Plomin 1993, Activity level, 20 mos., MZ	0.59 [−0.00, 1.18]
CTR_LTS, Plomin 1993, Activity level, 14 mos., MZ	0.50 [−0.10, 1.10]
CTR_LTS, Emde 1992, Activity, 14 mos., MZ	0.50 [−0.09, 1.09]
CTR_LTS, Saudino 1996, Activity level, 24 mos., MZ	0.26 [−0.28, 0.80]
CTR_LTS, Saudino 1996, Activity level, 20 mos., MZ	0.27 [−0.27, 0.81]
CTR_LTS, Saudino 1996, Activity level, 14 mos., MZ	0.35 [−0.18, 0.88]
CTR_LTS, Plomin 1993, Activity level, 14−20 mos., MZ	0.26 [−0.39, 0.91]
CTR_LTS, Plomin 1993, Activity level, 14−20 mos., MZ	0.42 [−0.21, 1.05]
CTR_LTS, Plomin 1993, Activity level, 20 mos., MZ	0.26 [−0.39, 0.91]
CTR_LTS, Plomin 1993, Activity level, 14 mos., MZ	0.42 [−0.19, 1.03]
CTR_LTS, Emde 1992, Activity, 14 mos., MZ	0.42 [−0.17, 1.01]
TCTH, Chen 1990, Activity level, 6 mos., MZ	0.66 [−0.01, 1.33]
Pooled_sample, Goldsmith 1999, Activity level, 3−16 mos., MZ	0.54 [−0.00, 1.08]
MTS, Saudino 1991, Activity level, 7 mos., MZ	0.82 [0.22, 1.42]
MTS, Saudino 1991, Activity level, 7 mos., MZ	0.76 [0.12, 1.40]
LTS, Riese 1990, Activity−sleep, 0−1 mos., MZ	0.22 [−0.53, 0.97]
LTS, Riese 1990, Activity−awake, 0−1 mos., MZ	0.09 [−0.67, 0.85]
LTS, Matheny 1976, Energy, 3−12 mos., MZ	0.44 [−0.25, 1.13]
LTS, Matheny 1976, Energy, 18−30 mos., MZ	0.81 [0.23, 1.39]
LTS, Matheny 1983, Activity level, 6−18 mos., MZ	0.27 [−0.43, 0.97]
LTS, Matheny 1983, Activity level, 12−24 mos., MZ	0.52 [−0.12, 1.16]
LTS, Matheny 1980, Activity level, 3 mos., MZ	0.30 [−0.35, 0.95]
LTS, Davis 2015, Activity level, 24 mos., MZ	0.33 [−0.17, 0.83]
LTS, Davis 2015, Activity level, 18 mos., MZ	0.44 [−0.06, 0.94]
LTS, Davis 2015, Activity level, 12 mos., MZ	0.32 [−0.18, 0.82]
LTS, Davis 2015, Activity level, 9 mos., MZ	0.35 [−0.18, 0.88]
LTS, Davis 2015, Activity level, 6 mos., MZ	0.43 [−0.08, 0.94]
JEP, Saudino 2008, Activity/Impulsivity, 24 mos., MZM	0.78 [0.37, 1.19]
JEP, Saudino 2008, Activity/Impulsivity, 24 mos., MZF	0.79 [0.37, 1.21]
GS, Smith 2017, Activity level, 3 mos., MZ	0.93 [0.70, 1.16]
USDP, Stevenson 1985, Activity level, 12−24 mos., MZM	0.21 [−0.71, 1.13]
USDP, Stevenson 1985, Activity level, 12−24 mos., MZF	−0.06 [−1.02, 0.90]
CPP, Goldsmith 1981, Vigorous activity vs. psychomotor passivity, 8 mos., MZ	0.57 [0.03, 1.11]
CPP, Goldsmith 1981, Activity level, 8 mos., MZ	0.41 [−0.16, 0.98]
BUTP, Saudino 2012, Activity level, 24 mos., MZ	0.82 [0.39, 1.25]
BUTP, Saudino 2012, Activity level, 24 mos., MZ	0.54 [0.02, 1.06]
BUTP, Saudino 2018, Activity level, 24 mos., MZ	0.73 [0.26, 1.20]
BUTP, Illot 2010, Activity level, 24 mos., MZ	0.75 [0.29, 1.21]
RE Model for MZ Subset	0.62 [0.47, 0.77]

Figure 21A (Continued)

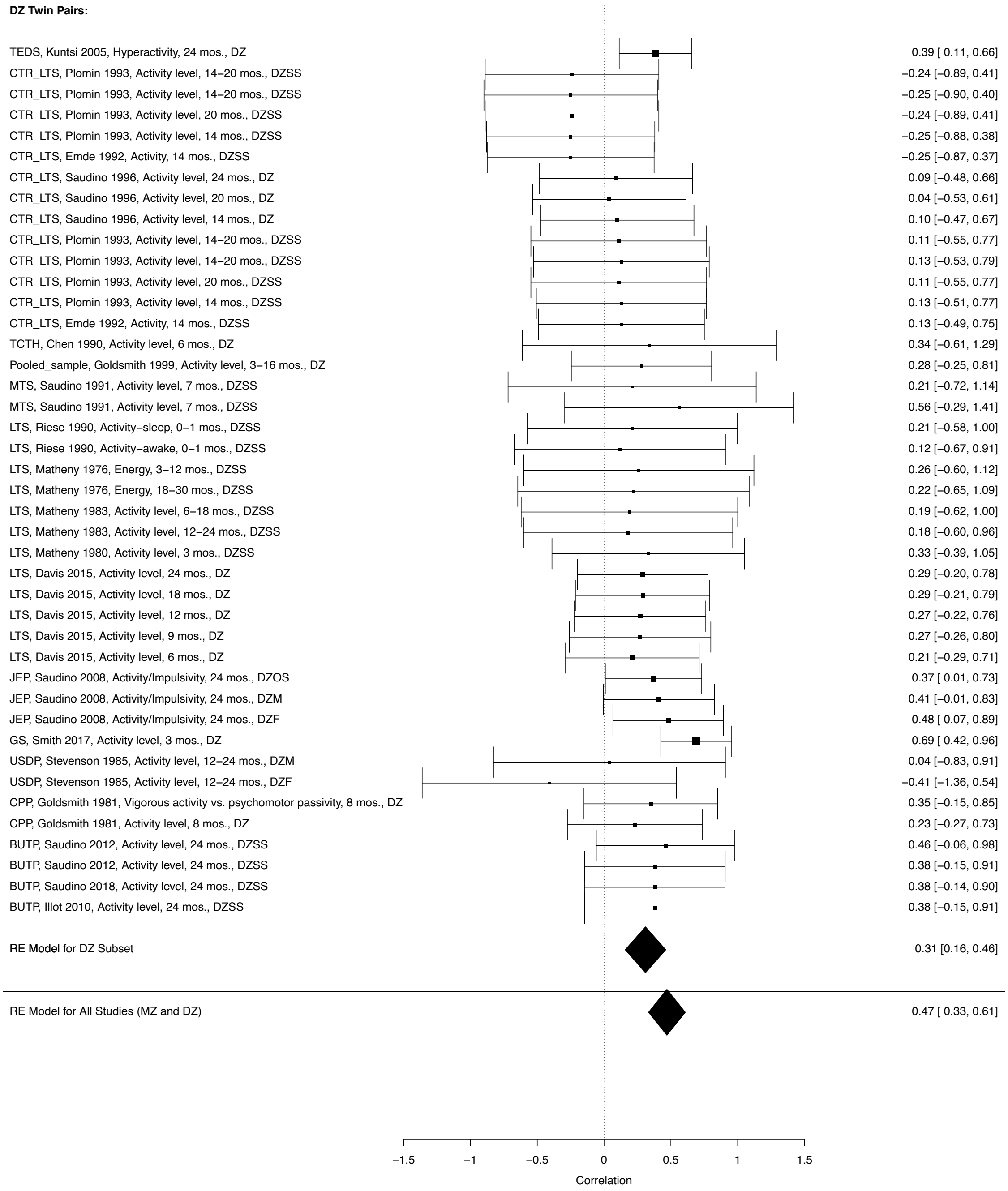


Figure 22A. Organization of Psychomotor Functions Forest Plot

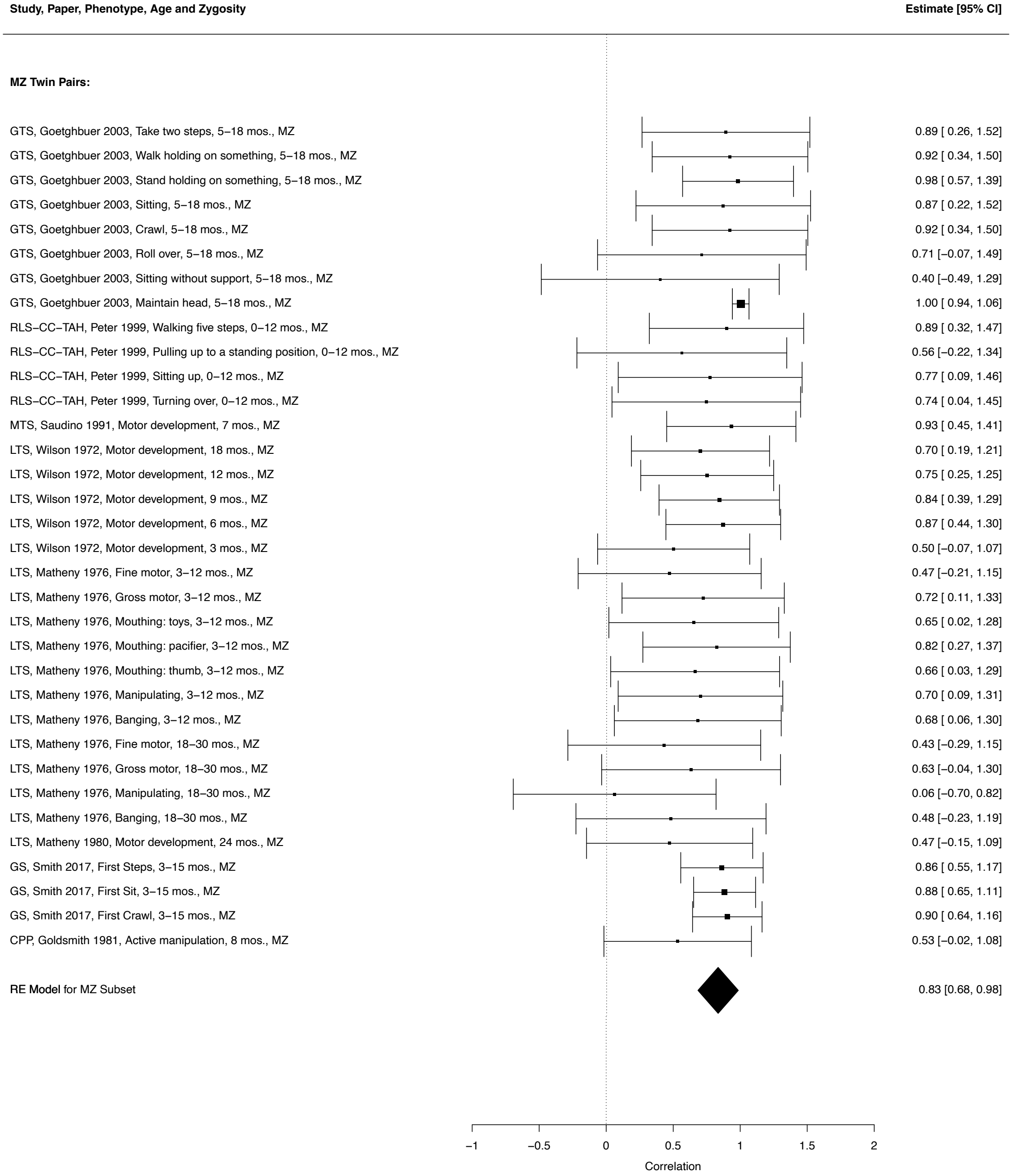


Figure 22A (Continued)

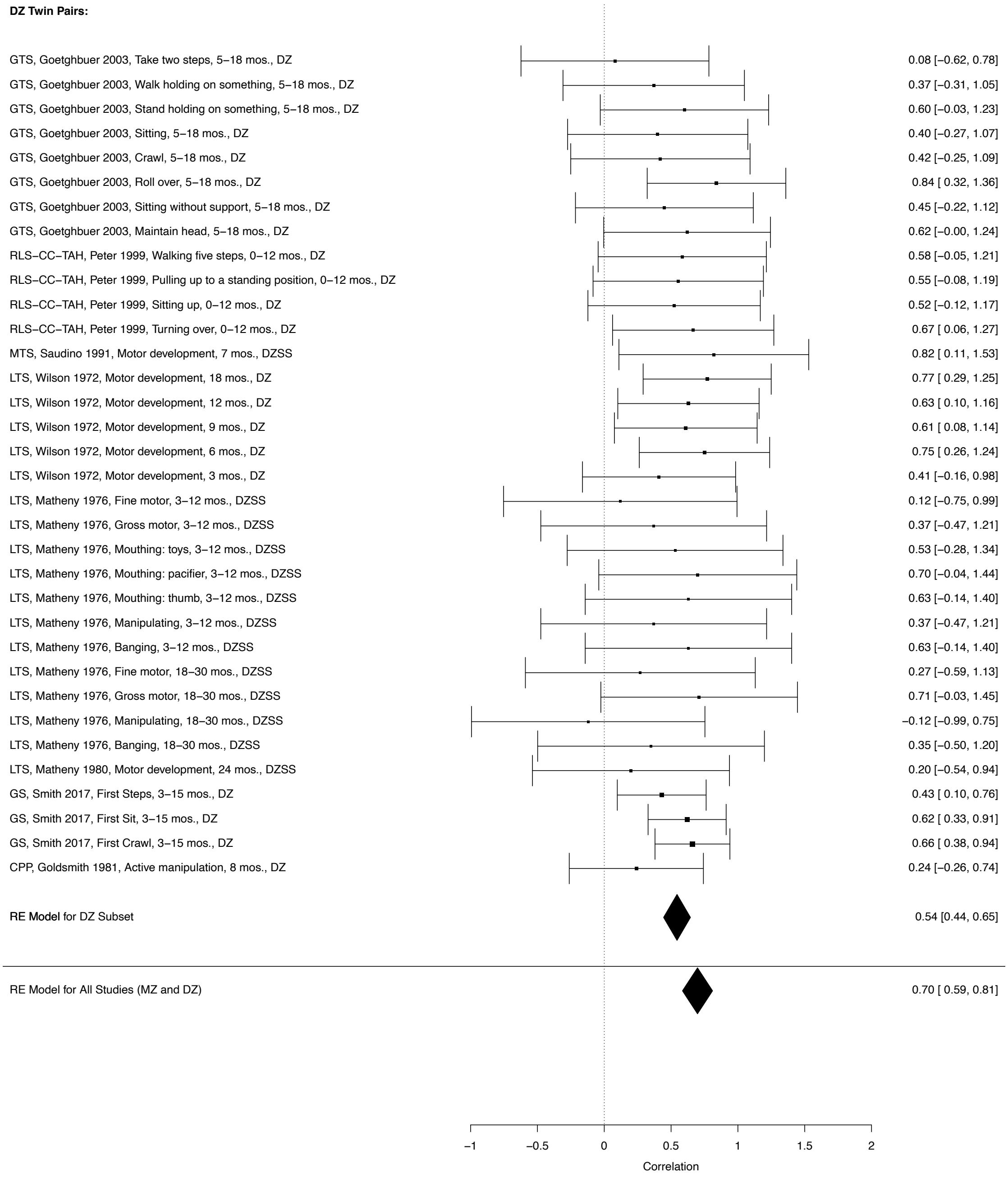


Figure 23A. Regulation of Emotion Forest Plot

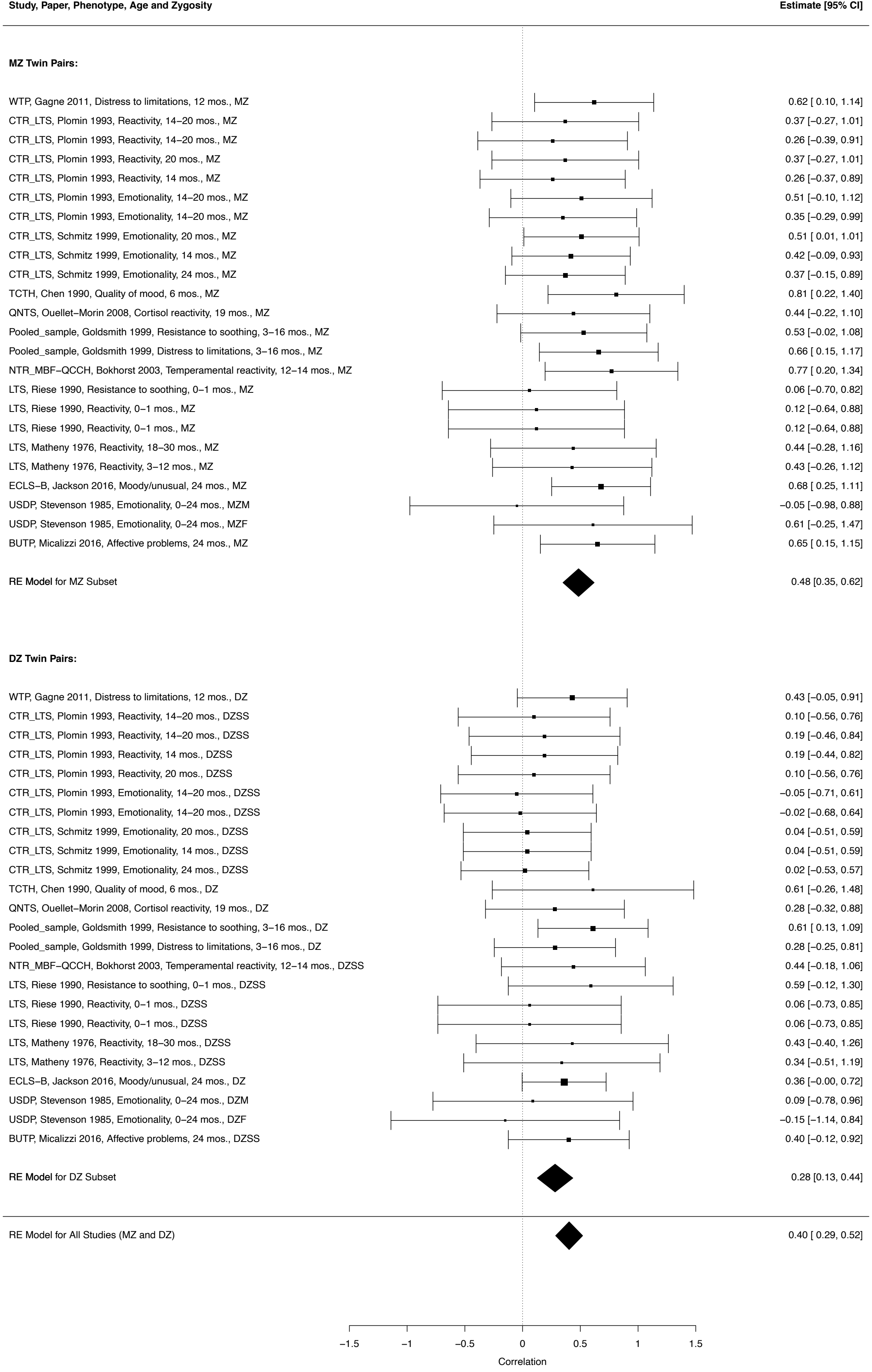
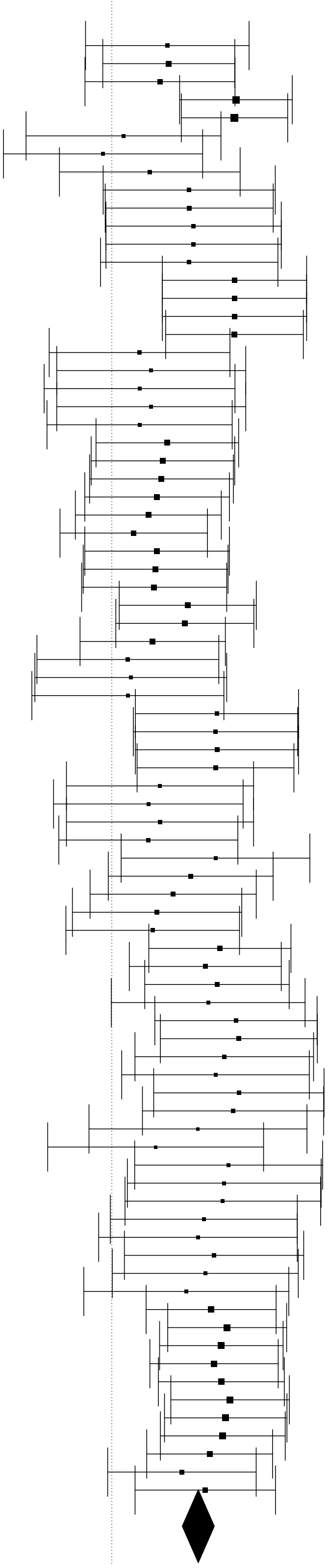


Figure 24A. Range of Emotion Forest Plot



Study, Paper, Phenotype, Age and Zygosity	Estimate [95% CI]
MZ Twin Pairs:	
WTP, Gagne 2011, Anger, 12 mos., MZ	0.38 [−0.18, 0.94]
WC_BTTYP, Planalp 2017, Positive affect, 12 mos., MZ	0.39 [−0.06, 0.84]
WC_BTTYP, Planalp 2017, Positive affect, 6 mos., MZ	0.33 [−0.18, 0.84]
WC_BTTYP, Planalp 2017, Smiling and laughter, 6 mos., MZ	0.85 [0.46, 1.24]
WC_BTTYP, Planalp 2017, Smiling and laughter, 12 mos., MZ	0.84 [0.48, 1.20]
CTR_TIP, DiLalla 1996, Negative affect, 7 mos., MZ	0.08 [−0.59, 0.75]
CTR_TIP, DiLalla 1996, Negative affect, 9 mos., MZ	−0.06 [−0.74, 0.62]
CTR_LTS, Emde 1992, Frustration, 14 mos., MZ	0.26 [−0.36, 0.88]
CTR_LTS, Plomin 1993, Positive hedonic tone, 14 mos., MZ	0.53 [−0.06, 1.12]
CTR_LTS, Emde 1992, Positive hedonic tone, 14 mos., MZ	0.53 [−0.05, 1.11]
CTR_LTS, Plomin 1993, Positive hedonic tone, 20 mos., MZ	0.56 [−0.04, 1.16]
CTR_LTS, Plomin 1993, Positive affect, 14−20 mos., MZ	0.56 [−0.04, 1.16]
CTR_LTS, Plomin 1993, Positive affect, 14−20 mos., MZ	0.53 [−0.08, 1.14]
CTR_LTS, Plomin 1993, Positive affect, 14−20 mos., MZ	0.84 [0.35, 1.33]
CTR_LTS, Plomin 1993, Positive affect, 14−20 mos., MZ	0.84 [0.35, 1.33]
CTR_LTS, Plomin 1993, Positive affect, 20 mos., MZ	0.84 [0.35, 1.33]
CTR_LTS, Plomin 1993, Positive affect, 14 mos., MZ	0.84 [0.37, 1.31]
CTR_LTS, Emde 1992, Overall mood, 14 mos., MZ	0.19 [−0.43, 0.81]
CTR_LTS, Plomin 1993, Affect, 14−20 mos., MZ	0.27 [−0.38, 0.92]
CTR_LTS, Plomin 1993, Affect, 14−20 mos., MZ	0.19 [−0.46, 0.84]
CTR_LTS, Plomin 1993, Overall mood, 20 mos., MZ	0.27 [−0.38, 0.92]
CTR_LTS, Plomin 1993, Affect, 14 mos., MZ	0.19 [−0.44, 0.82]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.38 [−0.11, 0.87]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.35 [−0.14, 0.84]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.34 [−0.15, 0.83]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.31 [−0.19, 0.81]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.25 [−0.25, 0.75]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.15 [−0.35, 0.65]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.31 [−0.19, 0.81]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.30 [−0.20, 0.80]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.29 [−0.21, 0.79]
CTR_LTS, Rhee 2012, Negative affect, 24 mos., MZ	0.52 [0.05, 0.99]
CTR_LTS, Rhee 2012, Negative affect, 20 mos., MZ	0.50 [0.03, 0.97]
CTR_LTS, Rhee 2012, Negative affect, 14 mos., MZ	0.28 [−0.22, 0.78]
CTR_LTS, Emde 1992, Negative hedonic tone, 14 mos., MZ	0.11 [−0.51, 0.73]
CTR_LTS, Plomin 1993, Negative affect, 14−20 mos., MZ	0.13 [−0.53, 0.79]
CTR_LTS, Plomin 1993, Negative affect, 14−20 mos., MZ	0.11 [−0.55, 0.77]
CTR_LTS, Plomin 1993, Negative affect, 14−20 mos., MZ	0.72 [0.16, 1.28]
CTR_LTS, Plomin 1993, Negative affect, 14−20 mos., MZ	0.71 [0.15, 1.27]
CTR_LTS, Plomin 1993, Negative affect, 20 mos., MZ	0.72 [0.16, 1.28]
CTR_LTS, Plomin 1993, Negative affect, 14 mos., MZ	0.71 [0.17, 1.25]
CTR_LTS, Plomin 1993, Affect, 14−20 mos., MZ	0.33 [−0.31, 0.97]
CTR_LTS, Plomin 1993, Affect, 14−20 mos., MZ	0.25 [−0.40, 0.90]
CTR_LTS, Plomin 1993, Affect, 20 mos., MZ	0.33 [−0.31, 0.97]
CTR_LTS, Emde 1992, Affect, 14 mos., MZ	0.25 [−0.36, 0.86]
TCTH, Chen 1990, Intensity of reaction, 6 mos., MZ	0.71 [0.06, 1.36]
QNTS, Schumann 2017, Negative affect, 18 mos., MZM	0.54 [−0.02, 1.10]
QNTS, Schumann 2017, Negative affect, 5 mos., MZF	0.42 [−0.15, 0.99]
QNTS, Schumann 2017, Negative affect, 18 mos., MZF	0.31 [−0.27, 0.89]
QNTS, Schumann 2017, Negative affect, 5 mos., MZM	0.28 [−0.31, 0.87]
Pooled_sample, Goldsmith 1999, Positive affect, 3−16 mos., MZ	0.74 [0.25, 1.23]
Pooled_sample, Goldsmith 1999, Negative affect, 3−16 mos., MZ	0.64 [0.12, 1.16]
Pooled_sample, Goldsmith 1999, Smiling and laughter, 3−16 mos., MZ	0.72 [0.23, 1.21]
MCTS, Stroganova 2000, Nonoriented discharges, 7−12 mos., MZ	0.66 [−0.00, 1.32]
MCTS, Stroganova 2000, Moderate tension, 7−12 mos., MZ	0.85 [0.29, 1.41]
MCTS, Stroganova 2000, Low tension, 7−12 mos., MZ	0.87 [0.33, 1.41]
MCTS, Stroganova 2000, High tension, 7−12 mos., MZ	0.77 [0.16, 1.38]
MCTS, Stroganova 2000, Fear, 7−12 mos., MZ	0.71 [0.07, 1.35]
LTS, Matheny 1989, Emotional tone, 24 mos., MZ	0.87 [0.29, 1.45]
LTS, Matheny 1989, Emotional tone, 18 mos., MZ	0.83 [0.21, 1.45]
LTS, Matheny 1989, Emotional tone, 12 mos., MZ	0.59 [−0.16, 1.34]
LTS, Riese 1990, Irritability, 0−1 mos., MZ	0.30 [−0.44, 1.04]
LTS, Matheny 1989, Fearfulness, 24 mos., MZ	0.80 [0.16, 1.44]
LTS, Matheny 1989, Fearfulness, 18 mos., MZ	0.77 [0.11, 1.43]
LTS, Matheny 1989, Fearfulness, 12 mos., MZ	0.76 [0.09, 1.43]
LTS, Matheny 1976, Tension, 3−12 mos., MZ	0.63 [−0.01, 1.27]
LTS, Matheny 1976, Tension, 18−30 mos., MZ	0.59 [−0.09, 1.27]
LTS, Matheny 1976, Fearfulness, 3−12 mos., MZ	0.70 [0.09, 1.31]
LTS, Matheny 1976, Emotional tone, 3−12 mos., MZ	0.64 [0.00, 1.28]
LTS, Matheny 1976, Emotional tone, 18−30 mos., MZ	0.51 [−0.19, 1.21]
JEP, Saudino 2008, Negative affect, 24 mos., MZM	0.68 [0.23, 1.13]
JEP, Saudino 2008, Internalizing, 24 mos., MZM	0.79 [0.38, 1.20]
JEP, Saudino 2008, General anxiety, 24 mos., MZM	0.75 [0.33, 1.17]
JEP, Saudino 2008, Depression withdrawal, 24 mos., MZM	0.70 [0.26, 1.14]
JEP, Saudino 2008, Negative affect, 24 mos., MZF	0.75 [0.32, 1.18]
JEP, Saudino 2008, Internalizing, 24 mos., MZF	0.81 [0.40, 1.22]
JEP, Saudino 2008, General anxiety, 24 mos., MZF	0.78 [0.36, 1.20]
JEP, Saudino 2008, Depression withdrawal, 24 mos., MZF	0.76 [0.33, 1.19]
ECLS−B, Jackson 2016, Demanding/angry, 24 mos., MZ	0.67 [0.24, 1.10]
ECLS−B, Roisman 2006, Fussiness and demanding behaviour, 9 mos., MZ	0.48 [−0.03, 0.99]
ECLS−B, Roisman 2006, Positive and negative affect, 9 mos., MZ	0.64 [0.16, 1.12]
RE Model for MZ Subset	0.59 [0.48, 0.70]

Figure 24A (Continued)

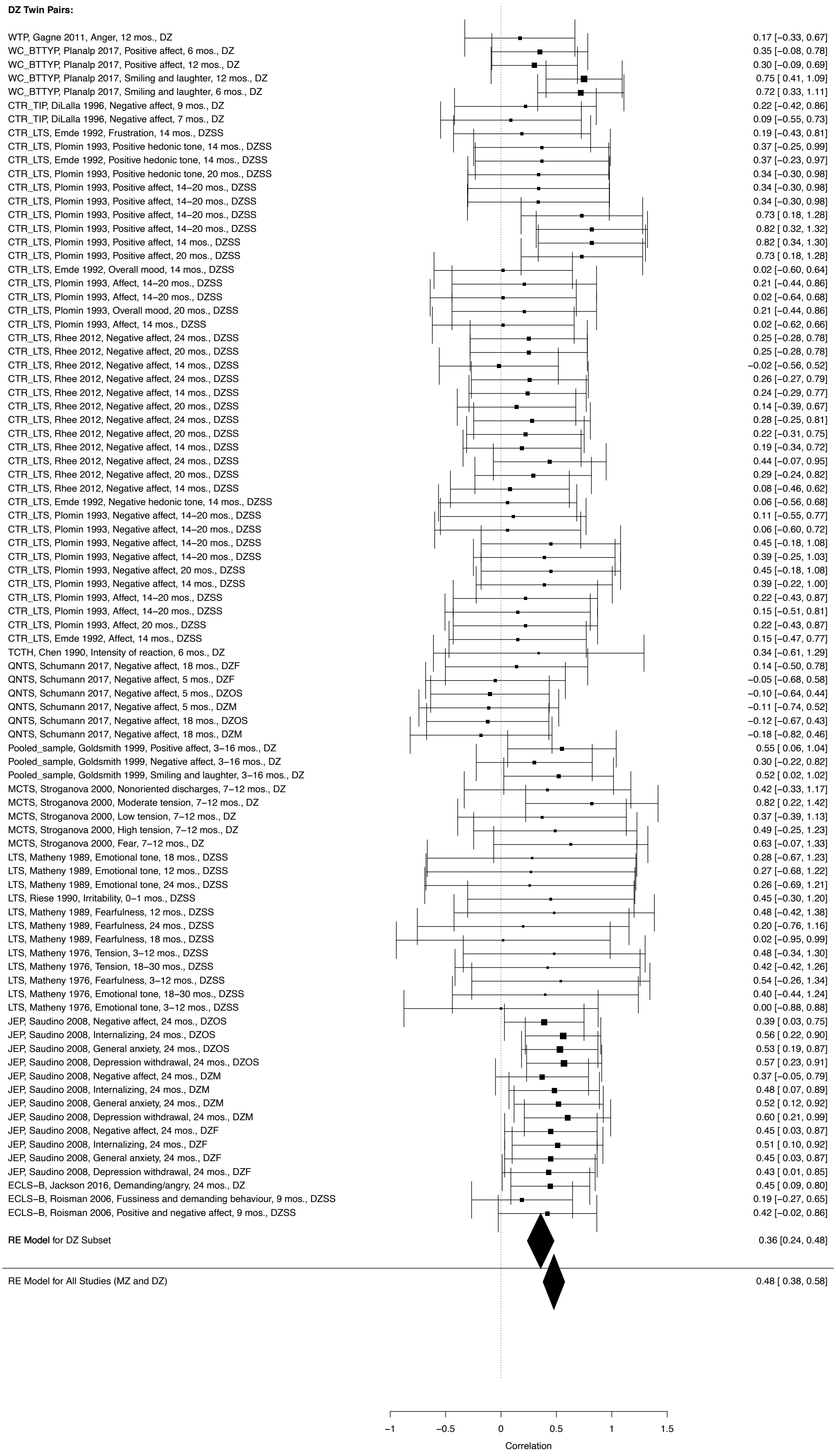


Figure 25A. Expression of Language Forest Plot

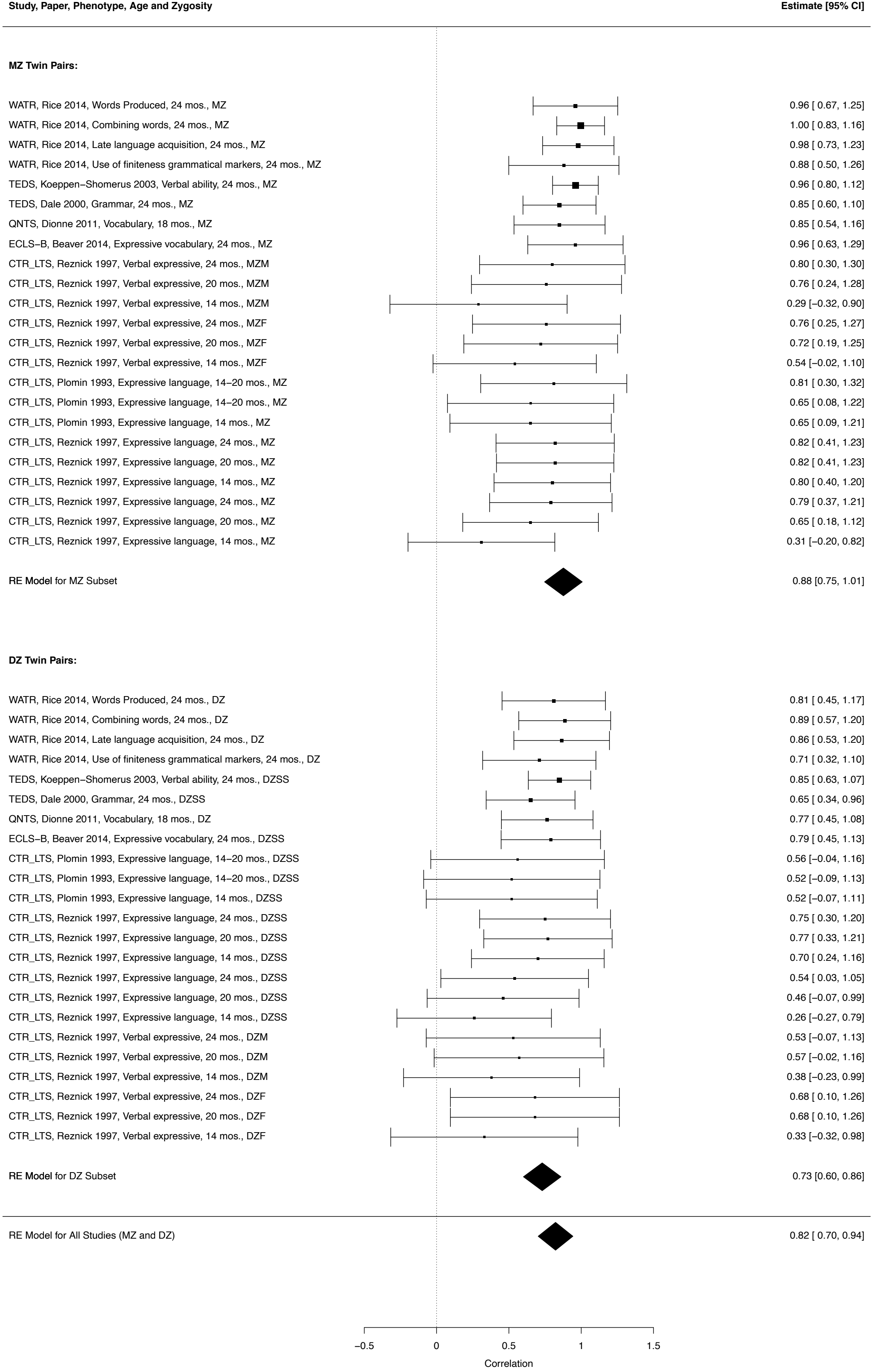
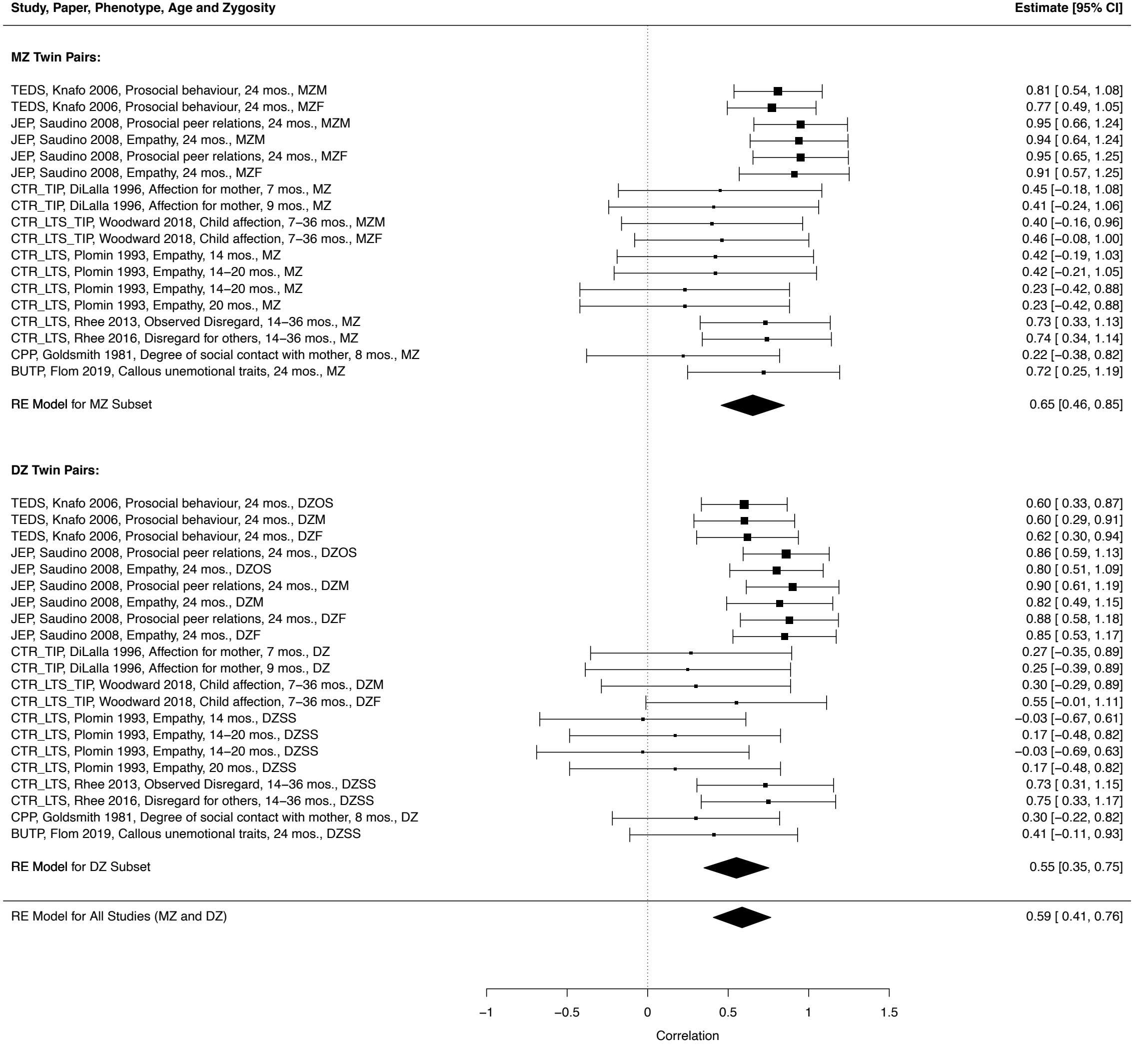


Figure 26A. Respect and Warmth in Relationships Forest Plot



Study, Paper, Phenotype, Age and Zygosity	Estimate [95% CI]
MZ Twin Pairs:	
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., MZM	0.81 [0.54, 1.08]
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., MZF	0.77 [0.49, 1.05]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., MZM	0.95 [0.66, 1.24]
JEP, Saudino 2008, Empathy, 24 mos., MZM	0.94 [0.64, 1.24]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., MZF	0.95 [0.65, 1.25]
JEP, Saudino 2008, Empathy, 24 mos., MZF	0.91 [0.57, 1.25]
CTR_TIP, DiLalla 1996, Affection for mother, 7 mos., MZ	0.45 [−0.18, 1.08]
CTR_TIP, DiLalla 1996, Affection for mother, 9 mos., MZ	0.41 [−0.24, 1.06]
CTR_LTS_TIP, Woodward 2018, Child affection, 7–36 mos., MZM	0.40 [−0.16, 0.96]
CTR_LTS_TIP, Woodward 2018, Child affection, 7–36 mos., MZF	0.46 [−0.08, 1.00]
CTR_LTS, Plomin 1993, Empathy, 14 mos., MZ	0.42 [−0.19, 1.03]
CTR_LTS, Plomin 1993, Empathy, 14–20 mos., MZ	0.42 [−0.21, 1.05]
CTR_LTS, Plomin 1993, Empathy, 14–20 mos., MZ	0.23 [−0.42, 0.88]
CTR_LTS, Plomin 1993, Empathy, 20 mos., MZ	0.23 [−0.42, 0.88]
CTR_LTS, Rhee 2013, Observed Disregard, 14–36 mos., MZ	0.73 [0.33, 1.13]
CTR_LTS, Rhee 2016, Disregard for others, 14–36 mos., MZ	0.74 [0.34, 1.14]
CPP, Goldsmith 1981, Degree of social contact with mother, 8 mos., MZ	0.22 [−0.38, 0.82]
BUTP, Flom 2019, Callous unemotional traits, 24 mos., MZ	0.72 [0.25, 1.19]
RE Model for MZ Subset	0.65 [0.46, 0.85]
DZ Twin Pairs:	
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., DZOS	0.60 [0.33, 0.87]
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., DZM	0.60 [0.29, 0.91]
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., DZF	0.62 [0.30, 0.94]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., DZOS	0.86 [0.59, 1.13]
JEP, Saudino 2008, Empathy, 24 mos., DZOS	0.80 [0.51, 1.09]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., DZM	0.90 [0.61, 1.19]
JEP, Saudino 2008, Empathy, 24 mos., DZM	0.82 [0.49, 1.15]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., DZF	0.88 [0.58, 1.18]
JEP, Saudino 2008, Empathy, 24 mos., DZF	0.85 [0.53, 1.17]
CTR_TIP, DiLalla 1996, Affection for mother, 7 mos., DZ	0.27 [−0.35, 0.89]
CTR_TIP, DiLalla 1996, Affection for mother, 9 mos., DZ	0.25 [−0.39, 0.89]
CTR_LTS_TIP, Woodward 2018, Child affection, 7–36 mos., DZM	0.30 [−0.29, 0.89]
CTR_LTS_TIP, Woodward 2018, Child affection, 7–36 mos., DZF	0.55 [−0.01, 1.11]
CTR_LTS, Plomin 1993, Empathy, 14 mos., DZSS	−0.03 [−0.67, 0.61]
CTR_LTS, Plomin 1993, Empathy, 14–20 mos., DZSS	0.17 [−0.48, 0.82]
CTR_LTS, Plomin 1993, Empathy, 14–20 mos., DZSS	−0.03 [−0.69, 0.63]
CTR_LTS, Plomin 1993, Empathy, 20 mos., DZSS	0.17 [−0.48, 0.82]
CTR_LTS, Rhee 2013, Observed Disregard, 14–36 mos., DZSS	0.73 [0.31, 1.15]
CTR_LTS, Rhee 2016, Disregard for others, 14–36 mos., DZSS	0.75 [0.33, 1.17]
CPP, Goldsmith 1981, Degree of social contact with mother, 8 mos., DZ	0.30 [−0.22, 0.82]
BUTP, Flom 2019, Callous unemotional traits, 24 mos., DZSS	0.41 [−0.11, 0.93]
RE Model for DZ Subset	0.55 [0.35, 0.75]
RE Model for All Studies (MZ and DZ)	0.59 [0.41, 0.76]

Figure 27A. Social Cues in Relationships Forest Plot

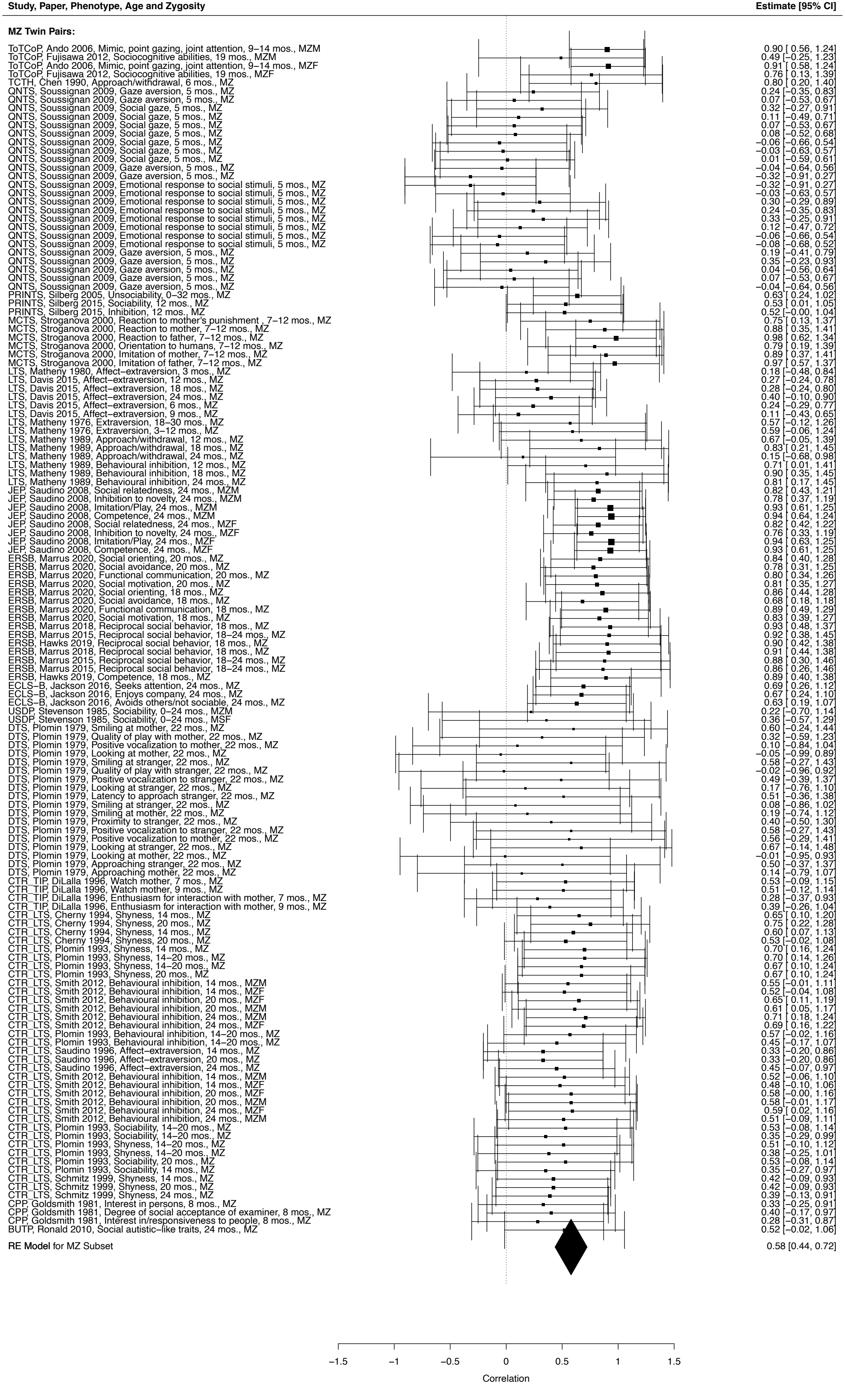


Figure 27A (Continued)

Figure 28A. Regulating Behaviors Within Interactions Forest Plot

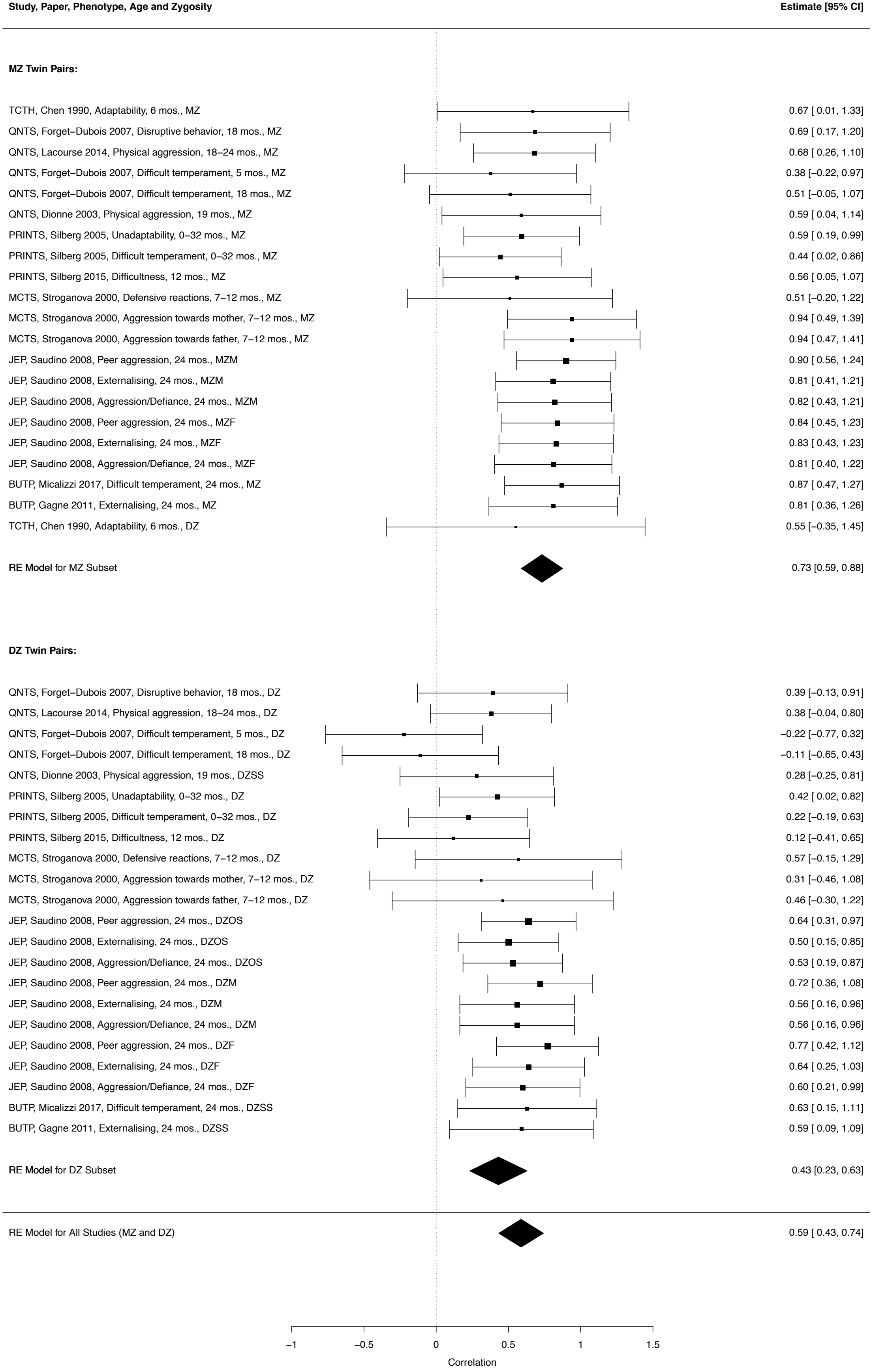


Figure 29A. Interacting According to Social Rules Forest Plot

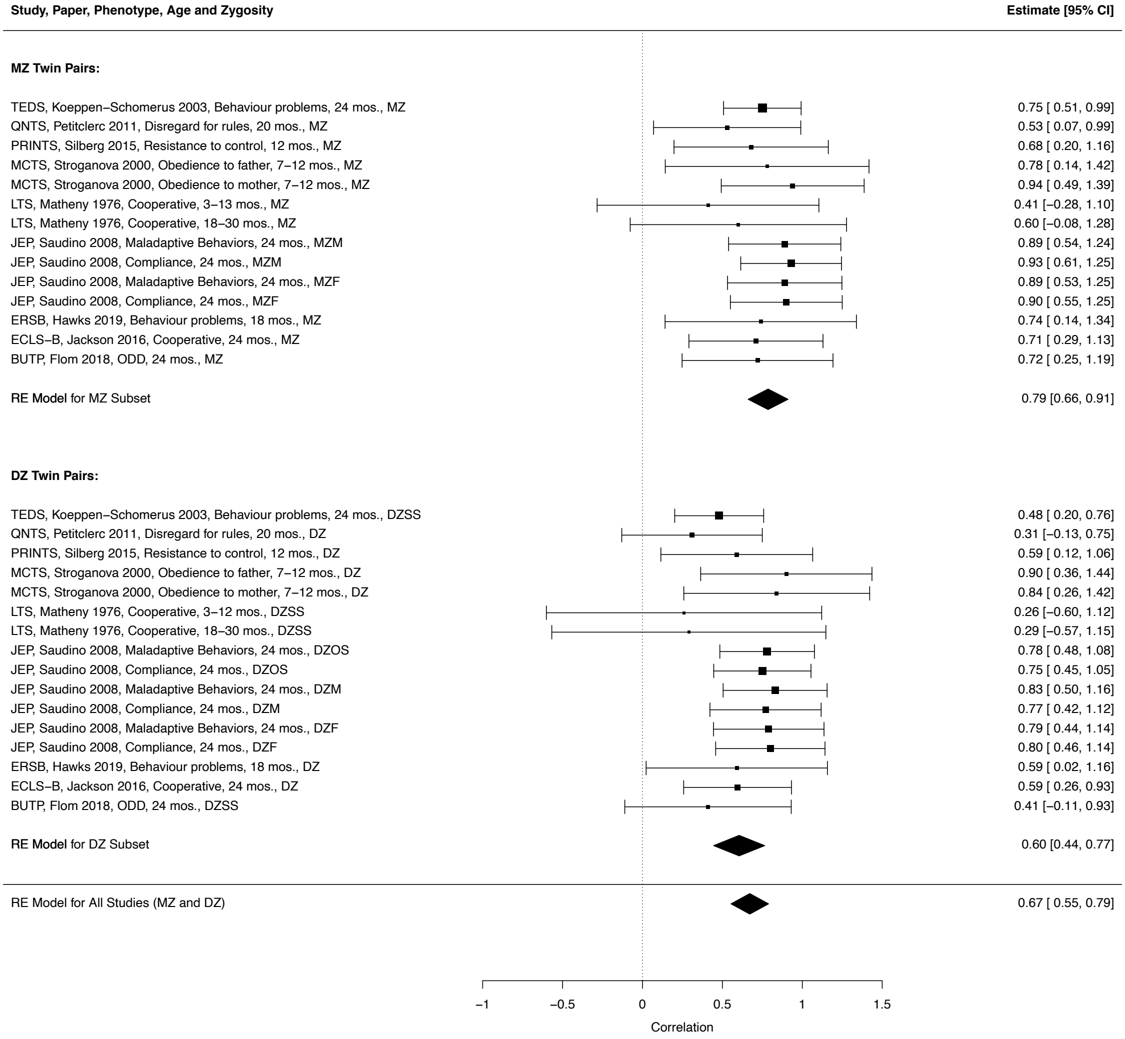


Figure 30A. Psychomotor Functions (Observer-Report) Forest Plot

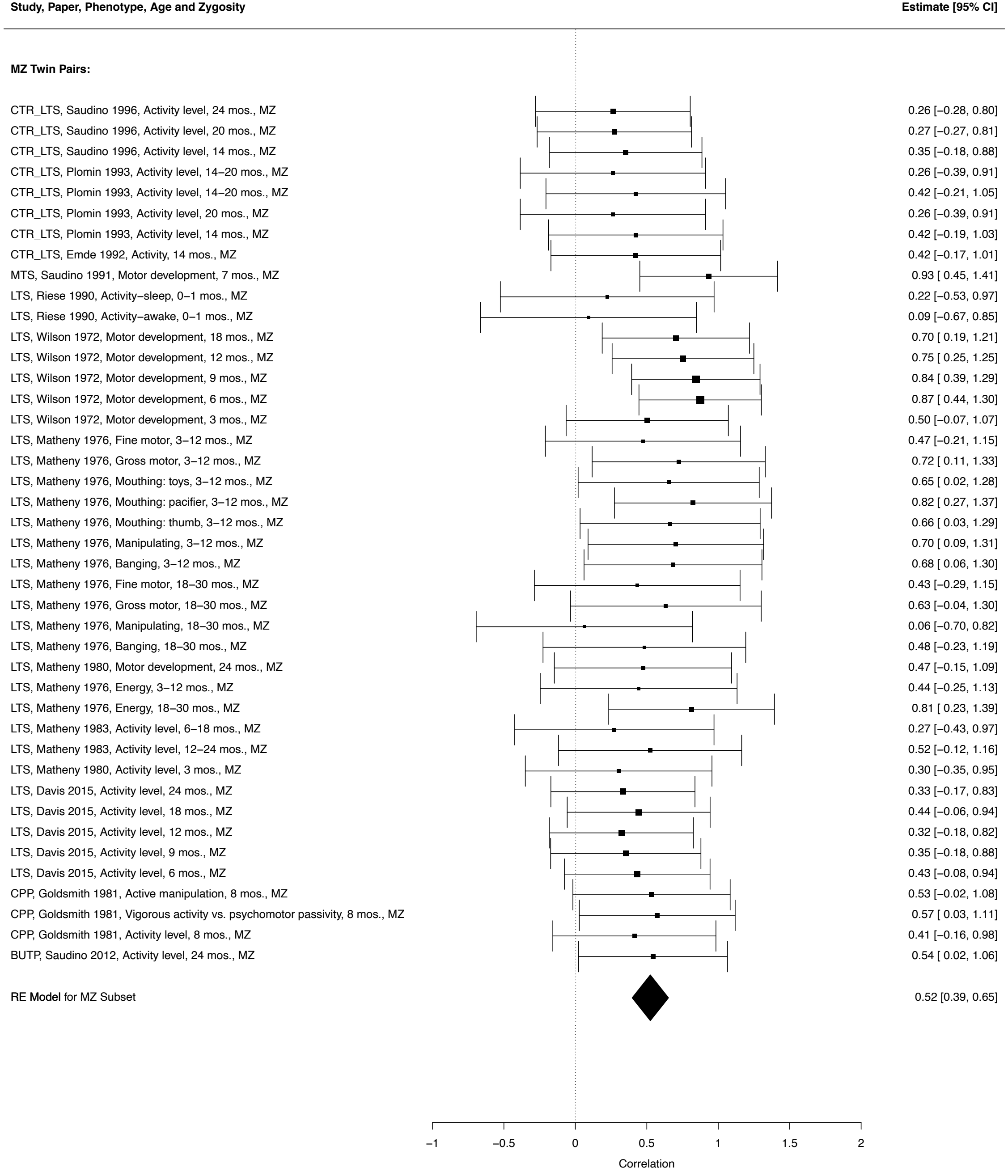


Figure 30A (Continued)

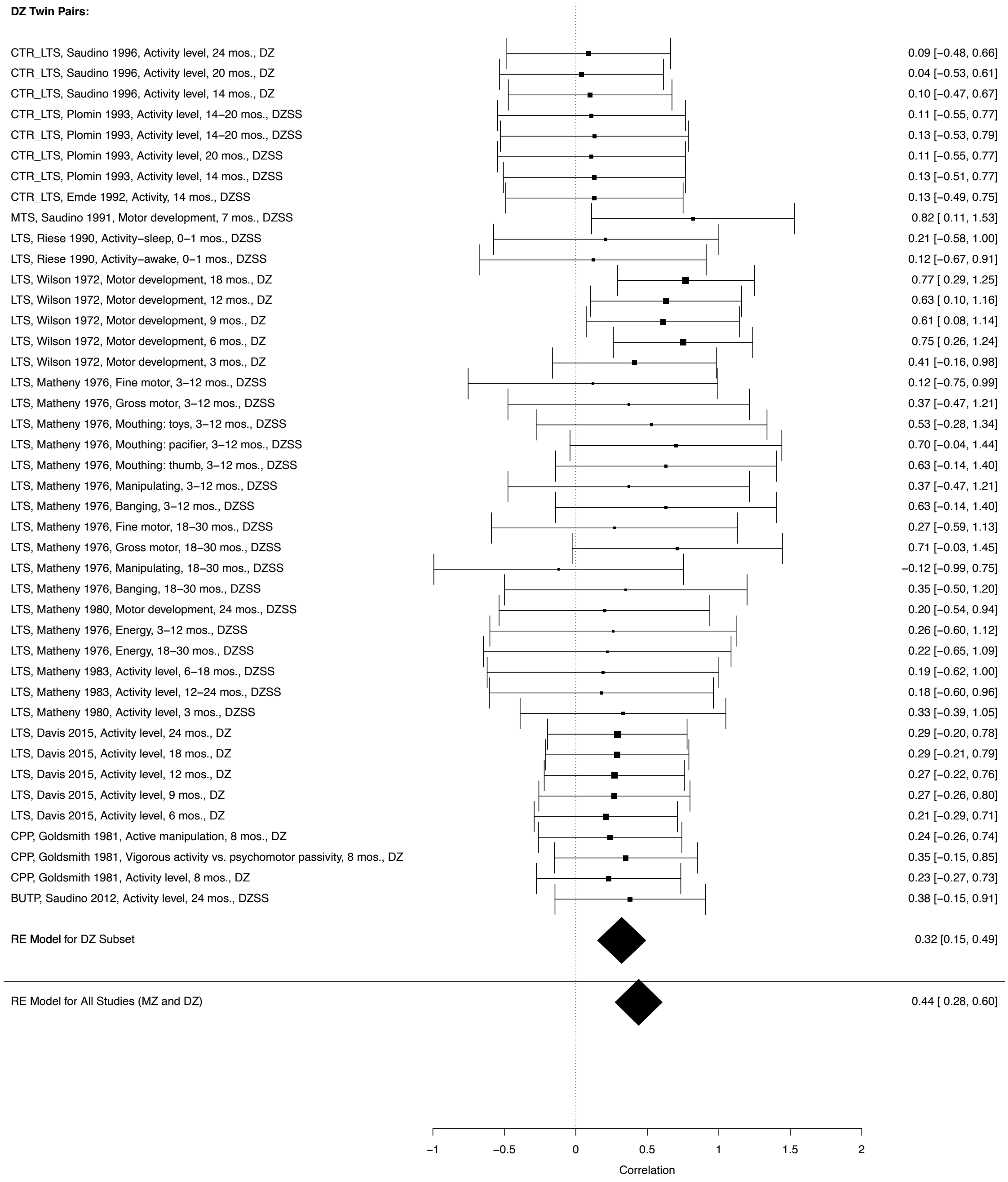


Figure 31A. Psychomotor Functions (Parent-Report) Forest Plot

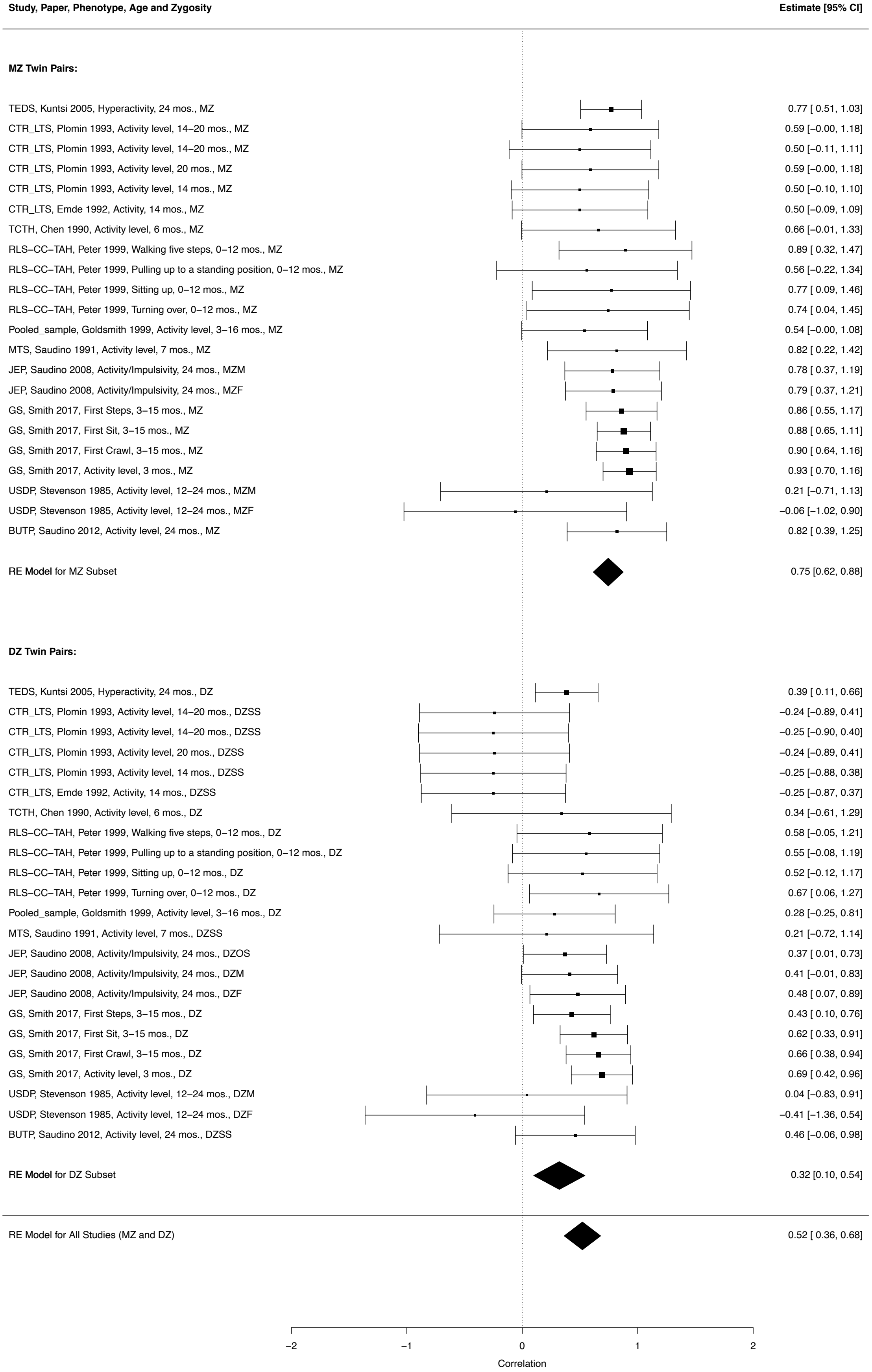


Figure 32A. Emotional Functions (Observer-Report) Forest Plot

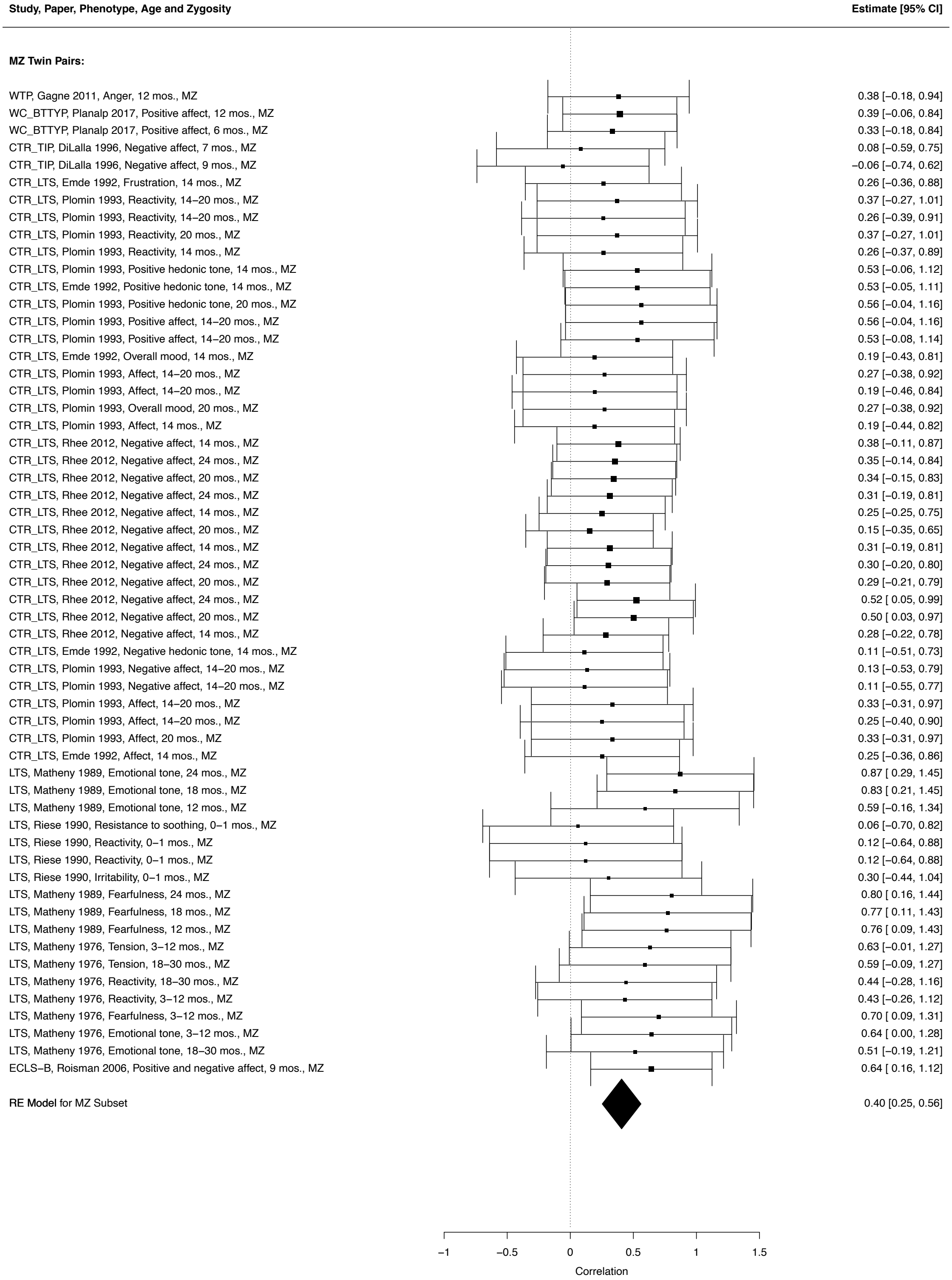


Figure 32A (Continued)

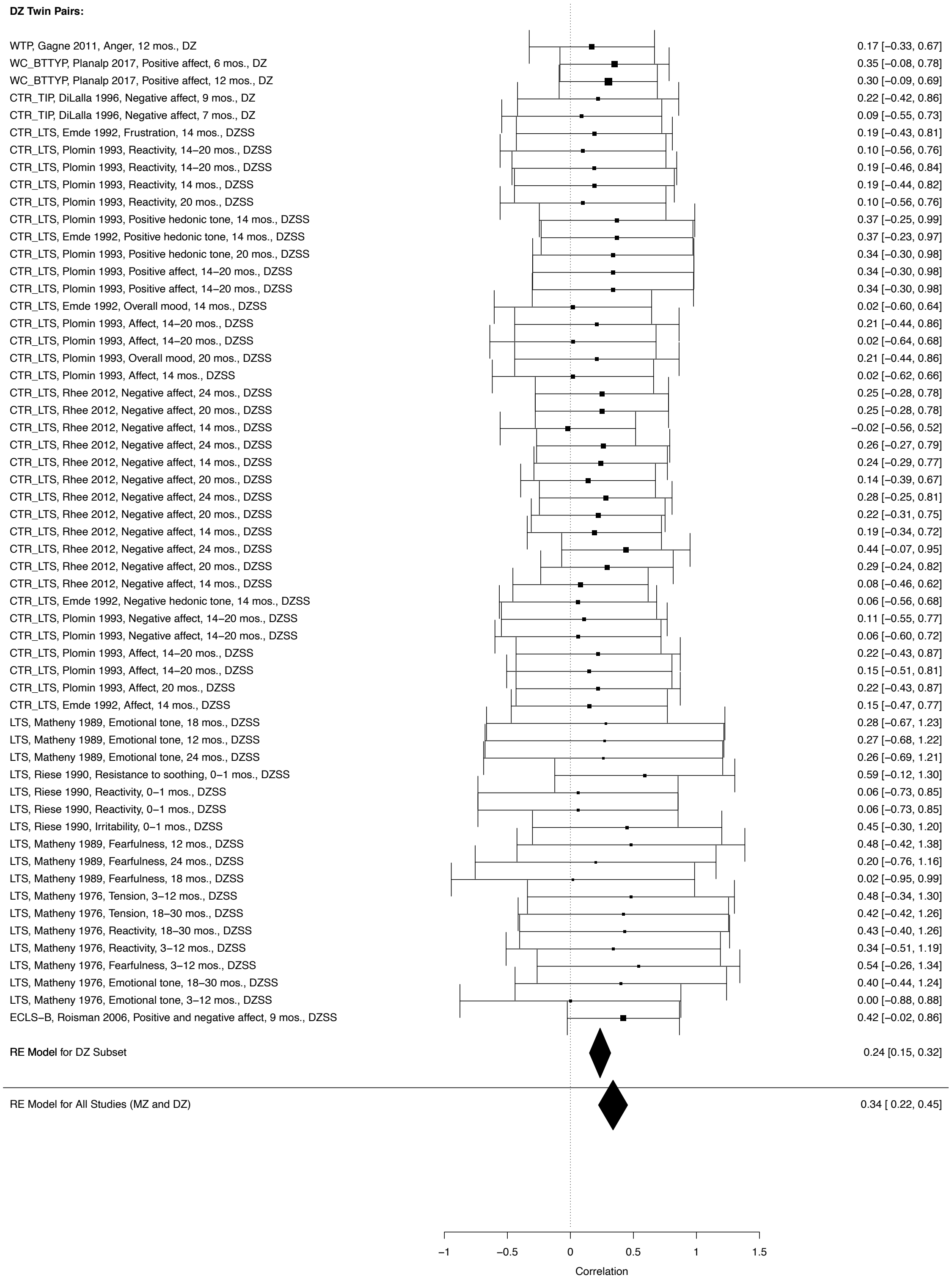


Figure 33A. Emotional Functions (Parent-Report) Forest Plot

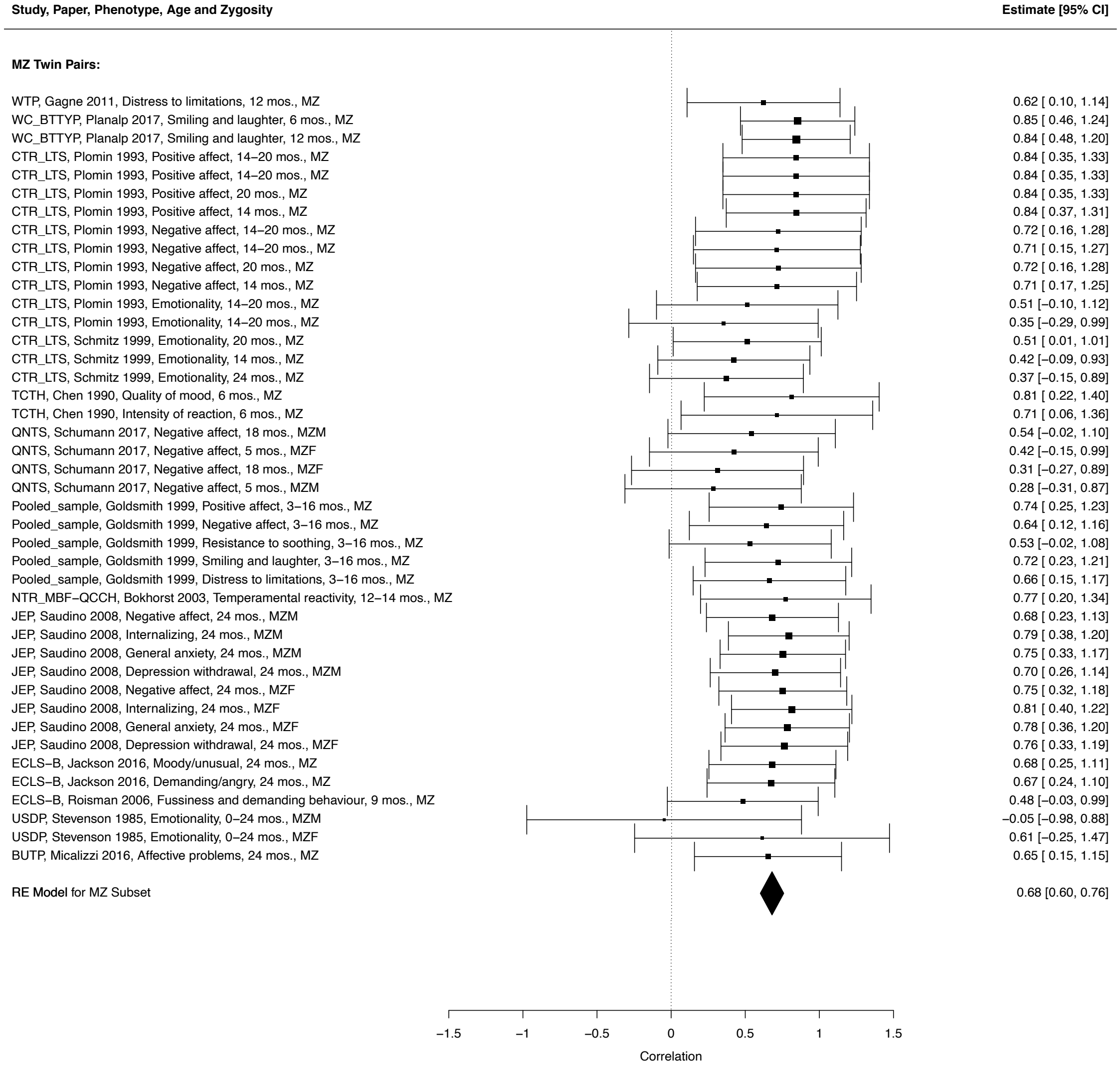


Figure 33A (Continued)

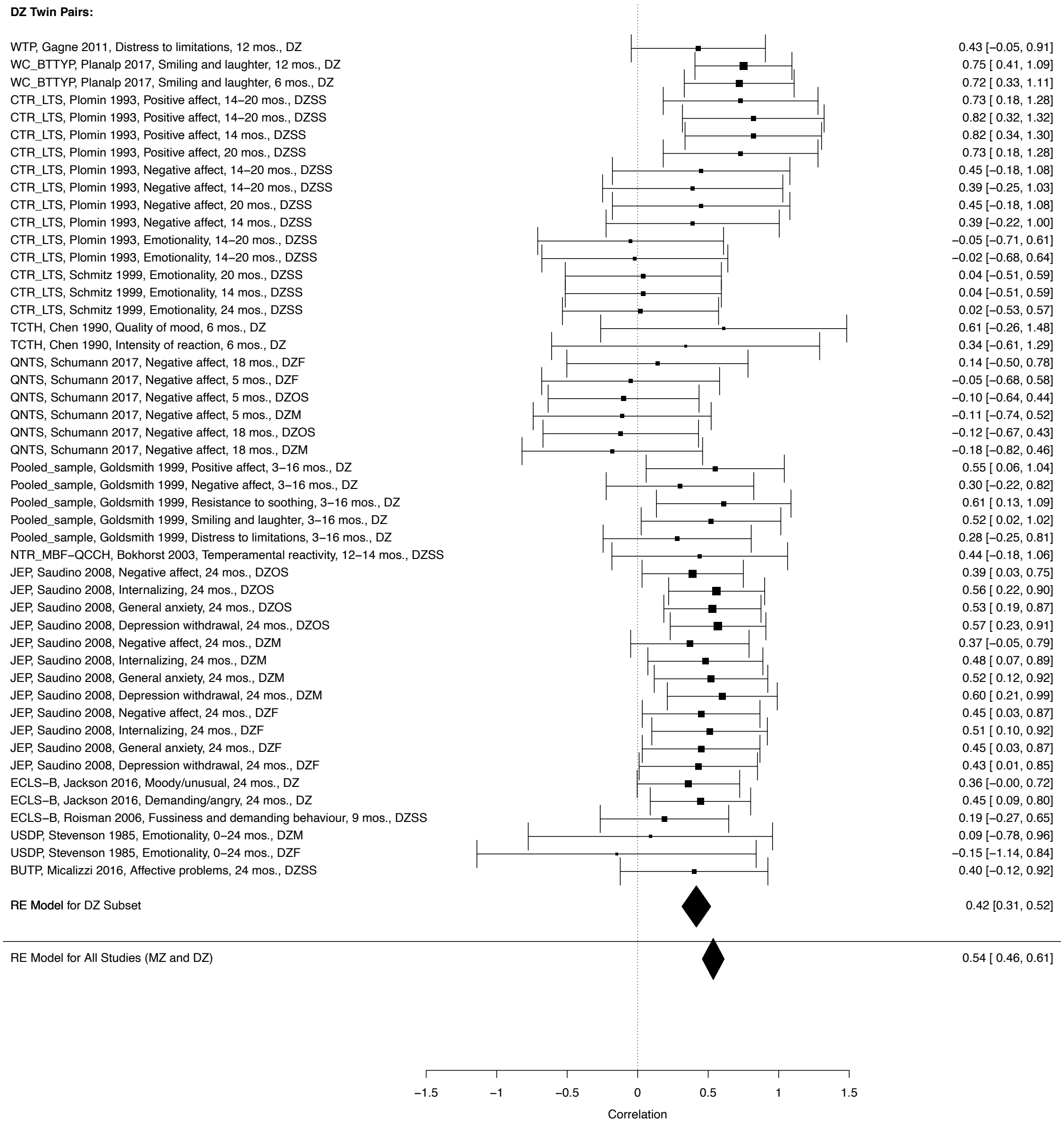
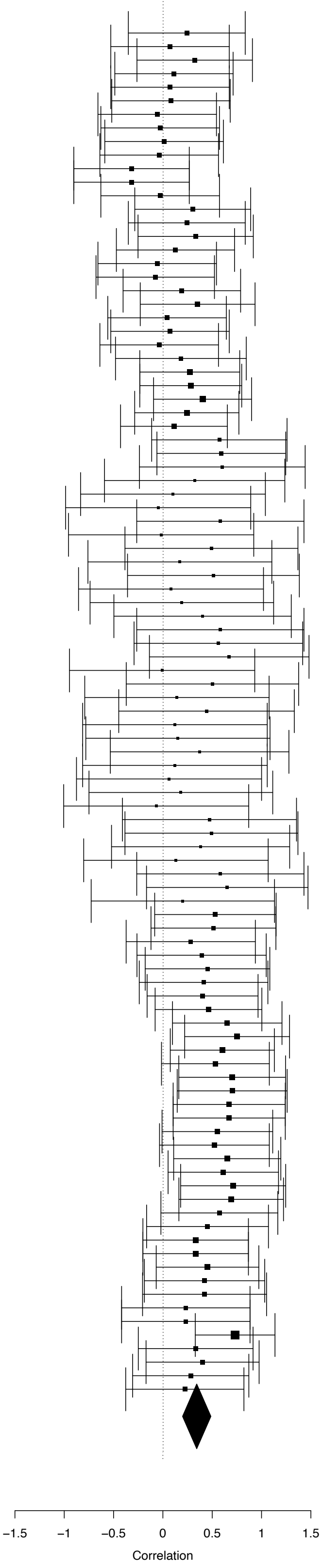


Figure 34A. Basic Interpersonal Interactions (Observer-Report) Forest Plot



Study, Paper, Phenotype, Age and Zygosity	Estimate [95% CI]
MZ Twin Pairs:	
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	0.24 [−0.35, 0.83]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	0.07 [−0.53, 0.67]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	0.32 [−0.27, 0.91]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	0.11 [−0.49, 0.71]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	0.07 [−0.53, 0.67]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	0.08 [−0.52, 0.68]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	−0.06 [−0.66, 0.54]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	−0.03 [−0.63, 0.57]
QNTS, Soussignan 2009, Social gaze, 5 mos., MZ	0.01 [−0.59, 0.61]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	−0.04 [−0.64, 0.56]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	−0.32 [−0.91, 0.27]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	−0.32 [−0.91, 0.27]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	−0.03 [−0.63, 0.57]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	0.30 [−0.29, 0.89]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	0.24 [−0.35, 0.83]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	0.33 [−0.25, 0.91]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	0.12 [−0.47, 0.72]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	−0.06 [−0.66, 0.54]
QNTS, Soussignan 2009, Emotional response to social stimuli, 5 mos., MZ	−0.08 [−0.68, 0.52]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	0.19 [−0.41, 0.79]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	0.35 [−0.23, 0.93]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	0.04 [−0.56, 0.64]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	0.07 [−0.53, 0.67]
QNTS, Soussignan 2009, Gaze aversion, 5 mos., MZ	−0.04 [−0.64, 0.56]
LTS, Matheny 1980, Affect−extraversion, 3 mos., MZ	0.18 [−0.48, 0.84]
LTS, Davis 2015, Affect−extraversion, 12 mos., MZ	0.27 [−0.24, 0.78]
LTS, Davis 2015, Affect−extraversion, 18 mos., MZ	0.28 [−0.24, 0.80]
LTS, Davis 2015, Affect−extraversion, 24 mos., MZ	0.40 [−0.10, 0.90]
LTS, Davis 2015, Affect−extraversion, 6 mos., MZ	0.24 [−0.29, 0.77]
LTS, Davis 2015, Affect−extraversion, 9 mos., MZ	0.11 [−0.43, 0.65]
LTS, Matheny 1976, Extraversion, 18–30 mos., MZ	0.57 [−0.12, 1.26]
LTS, Matheny 1976, Extraversion, 3–12 mos., MZ	0.59 [−0.06, 1.24]
DTS, Plomin 1979, Smiling at mother, 22 mos., MZ	0.60 [−0.24, 1.44]
DTS, Plomin 1979, Quality of play with mother, 22 mos., MZ	0.32 [−0.59, 1.23]
DTS, Plomin 1979, Positive vocalization to mother, 22 mos., MZ	0.10 [−0.84, 1.04]
DTS, Plomin 1979, Looking at mother, 22 mos., MZ	−0.05 [−0.99, 0.89]
DTS, Plomin 1979, Smiling at stranger, 22 mos., MZ	0.58 [−0.27, 1.43]
DTS, Plomin 1979, Quality of play with stranger, 22 mos., MZ	−0.02 [−0.96, 0.92]
DTS, Plomin 1979, Positive vocalization to stranger, 22 mos., MZ	0.49 [−0.39, 1.37]
DTS, Plomin 1979, Looking at stranger, 22 mos., MZ	0.17 [−0.76, 1.10]
DTS, Plomin 1979, Latency to approach stranger, 22 mos., MZ	0.51 [−0.36, 1.38]
DTS, Plomin 1979, Smiling at stranger, 22 mos., MZ	0.08 [−0.86, 1.02]
DTS, Plomin 1979, Smiling at mother, 22 mos., MZ	0.19 [−0.74, 1.12]
DTS, Plomin 1979, Proximity to stranger, 22 mos., MZ	0.40 [−0.50, 1.30]
DTS, Plomin 1979, Positive vocalization to stranger, 22 mos., MZ	0.58 [−0.27, 1.43]
DTS, Plomin 1979, Positive vocalization to mother, 22 mos., MZ	0.56 [−0.29, 1.41]
DTS, Plomin 1979, Looking at stranger, 22 mos., MZ	0.67 [−0.14, 1.48]
DTS, Plomin 1979, Looking at mother, 22 mos., MZ	−0.01 [−0.95, 0.93]
DTS, Plomin 1979, Approaching stranger, 22 mos., MZ	0.50 [−0.37, 1.37]
DTS, Plomin 1979, Approaching mother, 22 mos., MZ	0.14 [−0.79, 1.07]
DTS, Plomin 1979, Cuddliness with stranger, 22 mos., MZ	0.44 [−0.45, 1.33]
DTS, Plomin 1979, Difference of response between mother and stranger – cuddliness, 22 mos., MZ	0.12 [−0.82, 1.06]
DTS, Plomin 1979, Cuddliness with mother, 22 mos., MZ	0.15 [−0.78, 1.08]
DTS, Plomin 1979, Difference of response between mother and stranger – smiling, 22 mos., MZ	0.37 [−0.53, 1.27]
DTS, Plomin 1979, Difference of response between mother and stranger – quality of play, 22 mos., MZ	0.12 [−0.82, 1.06]
DTS, Plomin 1979, Difference of response between mother and stranger – positive vocalizations, 22 mos., MZ	0.06 [−0.88, 1.00]
DTS, Plomin 1979, Difference of response between mother and stranger – looking, 22 mos., MZ	0.18 [−0.75, 1.11]
DTS, Plomin 1979, Touching stranger, 22 mos., MZ	−0.07 [−1.01, 0.87]
DTS, Plomin 1979, Touching mother, 22 mos., MZ	0.47 [−0.41, 1.35]
DTS, Plomin 1979, Difference of response between mother and stranger – touches, 22 mos., MZ	0.49 [−0.39, 1.37]
DTS, Plomin 1979, Difference of response between mother and stranger – smiling, 22 mos., MZ	0.38 [−0.52, 1.28]
DTS, Plomin 1979, Difference of response between mother and stranger – proximity, 22 mos., MZ	0.13 [−0.80, 1.06]
DTS, Plomin 1979, Difference of response between mother and stranger – positive vocalizations, 22 mos., MZ	0.58 [−0.27, 1.43]
DTS, Plomin 1979, Difference of response between mother and stranger – looking, 22 mos., MZ	0.65 [−0.17, 1.47]
DTS, Plomin 1979, Difference of response between mother and stranger – approach, 22 mos., MZ	0.20 [−0.73, 1.13]
CTR_TIP, DiLalla 1996, Watch mother, 7 mos., MZ	0.53 [−0.09, 1.15]
CTR_TIP, DiLalla 1996, Watch mother, 9 mos., MZ	0.51 [−0.12, 1.14]
CTR_TIP, DiLalla 1996, Enthusiasm for interaction with mother, 7 mos., MZ	0.28 [−0.37, 0.93]
CTR_TIP, DiLalla 1996, Enthusiasm for interaction with mother, 9 mos., MZ	0.39 [−0.26, 1.04]
CTR_TIP, DiLalla 1996, Affection for mother, 7 mos., MZ	0.45 [−0.18, 1.08]
CTR_TIP, DiLalla 1996, Affection for mother, 9 mos., MZ	0.41 [−0.24, 1.06]
CTR_LTS_TIP, Woodward 2018, Child affection, 7–36 mos., MZM	0.40 [−0.16, 0.96]
CTR_LTS_TIP, Woodward 2018, Child affection, 7–36 mos., MZF	0.46 [−0.08, 1.00]
CTR_LTS, Cherny 1994, Shyness, 14 mos., MZ	0.65 [0.10, 1.20]
CTR_LTS, Cherny 1994, Shyness, 20 mos., MZ	0.75 [0.22, 1.28]
CTR_LTS, Cherny 1994, Shyness, 14 mos., MZ	0.60 [0.07, 1.13]
CTR_LTS, Cherny 1994, Shyness, 20 mos., MZ	0.53 [−0.02, 1.08]
CTR_LTS, Plomin 1993, Shyness, 14 mos., MZ	0.70 [0.16, 1.24]
CTR_LTS, Plomin 1993, Shyness, 14–20 mos., MZ	0.70 [0.14, 1.26]
CTR_LTS, Plomin 1993, Shyness, 14–20 mos., MZ	0.67 [0.10, 1.24]
CTR_LTS, Plomin 1993, Shyness, 20 mos., MZ	0.67 [0.10, 1.24]
CTR_LTS, Smith 2012, Behavioural inhibition, 14 mos., MZM	0.55 [−0.01, 1.11]
CTR_LTS, Smith 2012, Behavioural inhibition, 14 mos., MZF	0.52 [−0.04, 1.08]
CTR_LTS, Smith 2012, Behavioural inhibition, 20 mos., MZF	0.65 [0.11, 1.19]
CTR_LTS, Smith 2012, Behavioural inhibition, 20 mos., MZM	0.61 [0.05, 1.17]
CTR_LTS, Smith 2012, Behavioural inhibition, 24 mos., MZM	0.71 [0.18, 1.24]
CTR_LTS, Smith 2012, Behavioural inhibition, 24 mos., MZF	0.69 [0.16, 1.22]
CTR_LTS, Plomin 1993, Behavioural inhibition, 14–20 mos., MZ	0.57 [−0.02, 1.16]
CTR_LTS, Plomin 1993, Behavioural inhibition, 14–20 mos., MZ	0.45 [−0.17, 1.07]
CTR_LTS, Saudino 1996, Affect−extraversion, 14 mos., MZ	0.33 [−0.20, 0.86]
CTR_LTS, Saudino 1996, Affect−extraversion, 20 mos., MZ	0.33 [−0.20, 0.86]
CTR_LTS, Saudino 1996, Affect−extraversion, 24 mos., MZ	0.45 [−0.07, 0.97]
CTR_LTS, Plomin 1993, Empathy, 14 mos., MZ	0.42 [−0.19, 1.03]
CTR_LTS, Plomin 1993, Empathy, 14–20 mos., MZ	0.42 [−0.21, 1.05]
CTR_LTS, Plomin 1993, Empathy, 14–20 mos., MZ	0.23 [−0.42, 0.88]
CTR_LTS, Plomin 1993, Empathy, 20 mos., MZ	0.23 [−0.42, 0.88]
CTR_LTS, Rhee 2013, Observed Disregard, 14–36 mos., MZ	0.73 [0.33, 1.13]
CPP, Goldsmith 1981, Interest in persons, 8 mos., MZ	0.33 [−0.25, 0.91]
CPP, Goldsmith 1981, Degree of social acceptance of examiner, 8 mos., MZ	0.40 [−0.17, 0.97]
CPP, Goldsmith 1981, Interest in/responsiveness to people, 8 mos., MZ	0.28 [−0.31, 0.87]
CPP, Goldsmith 1981, Degree of social contact with mother, 8 mos., MZ	0.22 [−0.38, 0.82]
RE Model for MZ Subset	0.34 [0.20, 0.48]

Figure 34A (Continued)

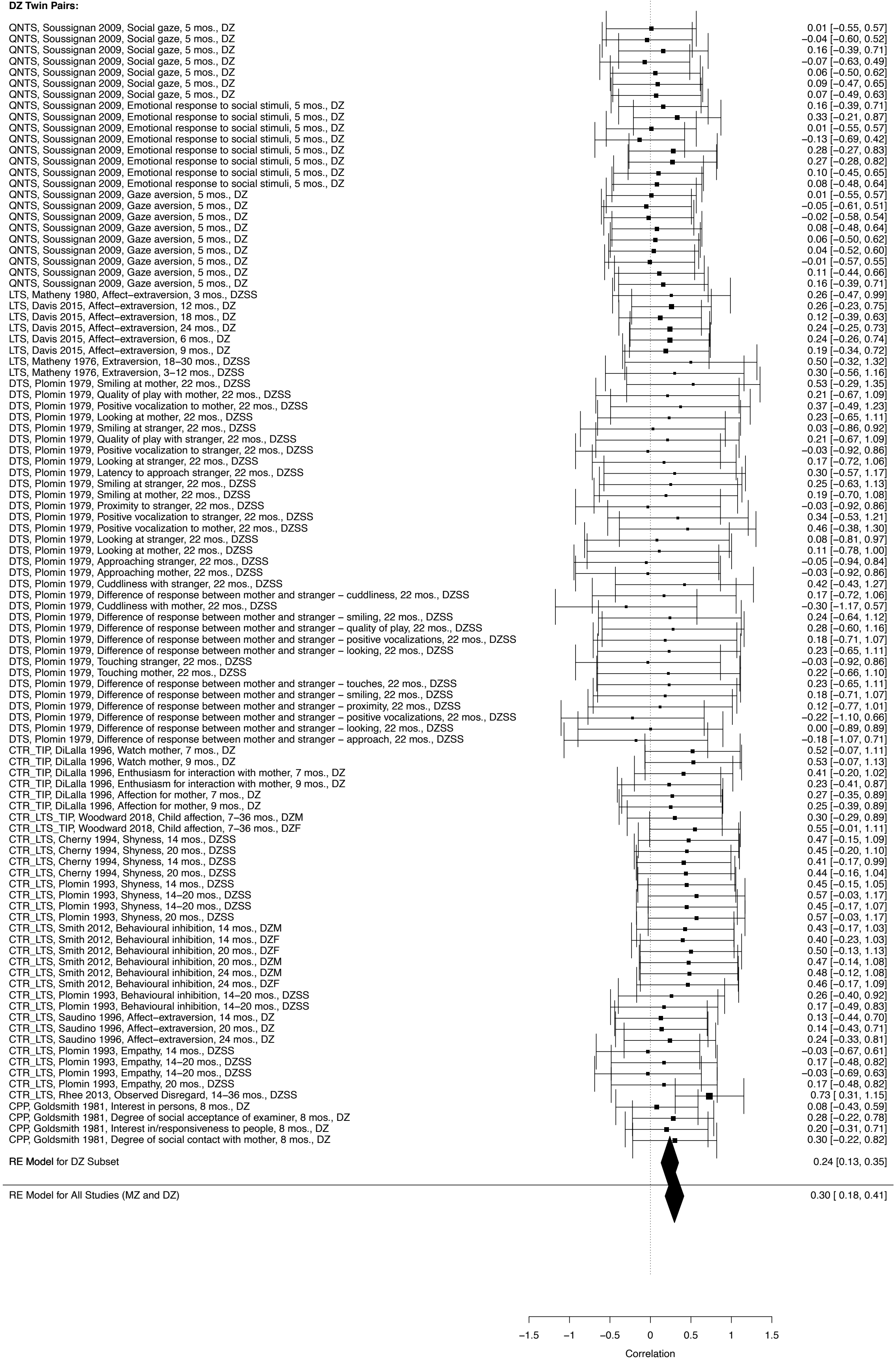
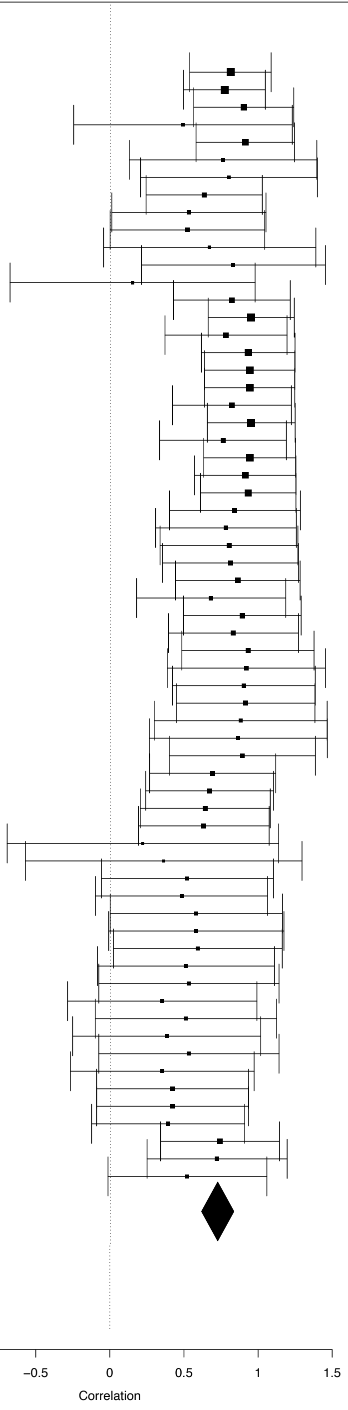
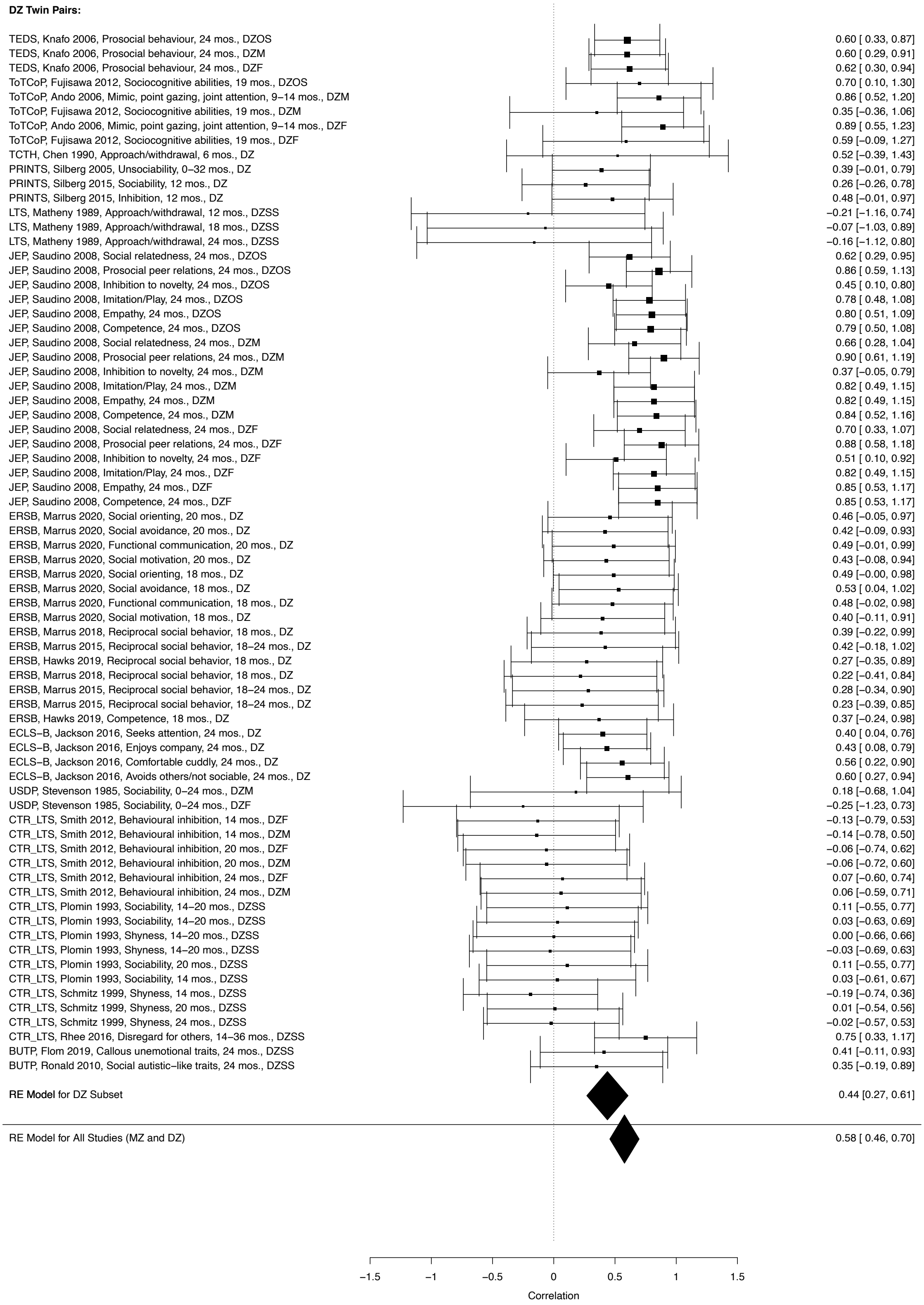


Figure 35A. Basic Interpersonal Interactions (Parent-Report) Forest Plot



Study, Paper, Phenotype, Age and Zygosity	**Estimate [95% CI]**
MZ Twin Pairs:	
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., MZM	0.81 [0.54, 1.08]
TEDS, Knafo 2006, Prosocial behaviour, 24 mos., MZF	0.77 [0.49, 1.05]
ToTCoP, Ando 2006, Mimic, point gazing, joint attention, 9–14 mos., MZM	0.90 [0.56, 1.24]
ToTCoP, Fujisawa 2012, Sociocognitive abilities, 19 mos., MZM	0.49 [−0.25, 1.23]
ToTCoP, Ando 2006, Mimic, point gazing, joint attention, 9–14 mos., MZF	0.91 [0.58, 1.24]
ToTCoP, Fujisawa 2012, Sociocognitive abilities, 19 mos., MZF	0.76 [0.13, 1.39]
TCTH, Chen 1990, Approach/withdrawal, 6 mos., MZ	0.80 [0.20, 1.40]
PRINTS, Silberg 2005, Unsociability, 0–32 mos., MZ	0.63 [0.24, 1.02]
PRINTS, Silberg 2015, Sociability, 12 mos., MZ	0.53 [0.01, 1.05]
PRINTS, Silberg 2015, Inhibition, 12 mos., MZ	0.52 [−0.00, 1.04]
LTS, Matheny 1989, Approach/withdrawal, 12 mos., MZ	0.67 [−0.05, 1.39]
LTS, Matheny 1989, Approach/withdrawal, 18 mos., MZ	0.83 [0.21, 1.45]
LTS, Matheny 1989, Approach/withdrawal, 24 mos., MZ	0.15 [−0.68, 0.98]
JEP, Saudino 2008, Social relatedness, 24 mos., MZM	0.82 [0.43, 1.21]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., MZM	0.95 [0.66, 1.24]
JEP, Saudino 2008, Inhibition to novelty, 24 mos., MZM	0.78 [0.37, 1.19]
JEP, Saudino 2008, Imitation/Play, 24 mos., MZM	0.93 [0.61, 1.25]
JEP, Saudino 2008, Empathy, 24 mos., MZM	0.94 [0.64, 1.24]
JEP, Saudino 2008, Competence, 24 mos., MZM	0.94 [0.64, 1.24]
JEP, Saudino 2008, Social relatedness, 24 mos., MZF	0.82 [0.42, 1.22]
JEP, Saudino 2008, Prosocial peer relations, 24 mos., MZF	0.95 [0.65, 1.25]
JEP, Saudino 2008, Inhibition to novelty, 24 mos., MZF	0.76 [0.33, 1.19]
JEP, Saudino 2008, Imitation/Play, 24 mos., MZF	0.94 [0.63, 1.25]
JEP, Saudino 2008, Empathy, 24 mos., MZF	0.91 [0.57, 1.25]
JEP, Saudino 2008, Competence, 24 mos., MZF	0.93 [0.61, 1.25]
ERSB, Marrus 2020, Social orienting, 20 mos., MZ	0.84 [0.40, 1.28]
ERSB, Marrus 2020, Social avoidance, 20 mos., MZ	0.78 [0.31, 1.25]
ERSB, Marrus 2020, Functional communication, 20 mos., MZ	0.80 [0.34, 1.26]
ERSB, Marrus 2020, Social motivation, 20 mos., MZ	0.81 [0.35, 1.27]
ERSB, Marrus 2020, Social orienting, 18 mos., MZ	0.86 [0.44, 1.28]
ERSB, Marrus 2020, Social avoidance, 18 mos., MZ	0.68 [0.18, 1.18]
ERSB, Marrus 2020, Functional communication, 18 mos., MZ	0.89 [0.49, 1.29]
ERSB, Marrus 2020, Social motivation, 18 mos., MZ	0.83 [0.39, 1.27]
ERSB, Marrus 2018, Reciprocal social behavior, 18 mos., MZ	0.93 [0.48, 1.37]
ERSB, Marrus 2015, Reciprocal social behavior, 18–24 mos., MZ	0.92 [0.38, 1.45]
ERSB, Hawks 2019, Reciprocal social behavior, 18 mos., MZ	0.90 [0.42, 1.38]
ERSB, Marrus 2018, Reciprocal social behavior, 18 mos., MZ	0.91 [0.44, 1.38]
ERSB, Marrus 2015, Reciprocal social behavior, 18–24 mos., MZ	0.88 [0.30, 1.46]
ERSB, Marrus 2015, Reciprocal social behavior, 18–24 mos., MZ	0.86 [0.26, 1.46]
ERSB, Hawks 2019, Competence, 18 mos., MZ	0.89 [0.40, 1.38]
ECLS–B, Jackson 2016, Seeks attention, 24 mos., MZ	0.69 [0.26, 1.12]
ECLS–B, Jackson 2016, Enjoys company, 24 mos., MZ	0.67 [0.24, 1.10]
ECLS–B, Jackson 2016, Comfortable cuddly, 24 mos., MZ	0.64 [0.20, 1.08]
ECLS–B, Jackson 2016, Avoids others/not sociable, 24 mos., MZ	0.63 [0.19, 1.07]
USDP, Stevenson 1985, Sociability, 0–24 mos., MZM	0.22 [−0.70, 1.14]
USDP, Stevenson 1985, Sociability, 0–24 mos., MZF	0.36 [−0.57, 1.29]
CTR_LTS, Smith 2012, Behavioural inhibition, 14 mos., MZM	0.52 [−0.06, 1.10]
CTR_LTS, Smith 2012, Behavioural inhibition, 14 mos., MZF	0.48 [−0.10, 1.06]
CTR_LTS, Smith 2012, Behavioural inhibition, 20 mos., MZF	0.58 [−0.00, 1.16]
CTR_LTS, Smith 2012, Behavioural inhibition, 20 mos., MZM	0.58 [−0.01, 1.17]
CTR_LTS, Smith 2012, Behavioural inhibition, 24 mos., MZF	0.59 [0.02, 1.16]
CTR_LTS, Smith 2012, Behavioural inhibition, 24 mos., MZM	0.51 [−0.09, 1.11]
CTR_LTS, Plomin 1993, Sociability, 14–20 mos., MZ	0.53 [−0.08, 1.14]
CTR_LTS, Plomin 1993, Sociability, 14–20 mos., MZ	0.35 [−0.29, 0.99]
CTR_LTS, Plomin 1993, Shyness, 14–20 mos., MZ	0.51 [−0.10, 1.12]
CTR_LTS, Plomin 1993, Shyness, 14–20 mos., MZ	0.38 [−0.25, 1.01]
CTR_LTS, Plomin 1993, Sociability, 20 mos., MZ	0.53 [−0.08, 1.14]
CTR_LTS, Plomin 1993, Sociability, 14 mos., MZ	0.35 [−0.27, 0.97]
CTR_LTS, Schmitz 1999, Shyness, 14 mos., MZ	0.42 [−0.09, 0.93]
CTR_LTS, Schmitz 1999, Shyness, 20 mos., MZ	0.42 [−0.09, 0.93]
CTR_LTS, Schmitz 1999, Shyness, 24 mos., MZ	0.39 [−0.13, 0.91]
CTR_LTS, Rhee 2016, Disregard for others, 14–36 mos., MZ	0.74 [0.34, 1.14]
BUTP, Flom 2019, Callous unemotional traits, 24 mos., MZ	0.72 [0.25, 1.19]
BUTP, Ronald 2010, Social autistic–like traits, 24 mos., MZ	0.52 [−0.02, 1.06]
RE Model for MZ Subset	0.72 [0.62, 0.83]

Figure 35A (Continued)



References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. University of Vermont, Research Center for Children, Youth & Families.

Adelantado-Renau, M., Moliner-Urdiales, D., Cavero-Redondo, I., Beltran-Valls, M. R., Martínez-Vizcaíno, V., & Álvarez-Bueno, C. (2019). Association between screen media use and academic performance among children and adolescents: A systematic review and meta-analysis. *JAMA Pediatrics, 173*(11), 1058-1067. https://doi.org/10.1001/jamapediatrics.2019.3176

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

Akerman, B. A., & Fischbein, S. (1992). Within-pair similarity in MZ and DZ twins from birth to eighteen years of age. *Acta Geneticae Medicae et Gemellologiae, 41*(2-3), 155-164. https://doi.org/10.1017/s0001566000002361

Allegrini, A. G., Selzam, S., Rimfeld, K., von Stumm, S., Pingault, J. B., & Plomin, R. (2019). Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry, 24*, 819–827. https://doi.org/10.1038/s41380-019-0394-4

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *The Journal of Abnormal Psychology, 112*, 545-557. https://doi.org/10.1037/0021-843X.112.4.545

Ando, J., Fujisawa, K. K., Hiraishi, K., Shikishima, C., Kawamoto, T., Nozaki, M., Yamagata, S., Takahashi, Y., Suzuki, K., Someya, Y., Ozaki, K., Deno, M., Tanaka, M., Sasaki, S., Toda, T., Kobayashi, K., Sakagami, M., Okada, M., Kijima, N., Takizawa, R., & Murayama, K. (2019). Psychosocial twin cohort studies in Japan: The Keio Twin Research Center (KoTReC). *Twin Research and Human Genetics, 22*(6), 591-596. https://doi.org/10.1017/thg.2019.109

Ando, J., Nonaka, K., Ozaki, K., Sato, N., Fujisawa, K. K., Suzuki, K., Yamagata, S., Takahashi, Y., Nakajima, R., Kato, N., & Ooki, S. (2006). The Tokyo Twin Cohort Project: overview and initial findings. *Twin Research and Human Genetics, 9*(6), 817-826. https://doi.org/10.1375/183242706779462480

Armstrong-Carter, E., Trejo, S., Hill, L. J. B., Crossley, K. L., Mason, D., & Domingue, B. W. (2020). The earliest origins of genetic nurture: The prenatal environment mediates the association between maternal genetics and child development. *Psychological Science, 31*(7), 781-791. https://doi.org/10.1177/0956797620917209

Arnett, J. J. (2008, Oct). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602-614. https://doi.org/10.1037/0003-066x.63.7.602

Asbury, K., & Plomin, R. (2013). *G is for genes: The impact of genetics on education and achievement*. John Wiley & Sons.

Austerberry, C., & Fearon, P. (2021). An overview of developmental behavioral genetics. In L. Provenzi & R. Montirosso (Eds.), *Developmental human behavioral epigenetics* (Vol. 23, pp. 59-80). Academic Press. https://doi.org/10.1016/B978-0-12-819262-7.00003-9

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., & Reiss, D. (2021, March 24). Evocative effects of genetic influences underlying the development of intellectual ability on the early caregiving environment. *OSF Preregistration*. https://doi.org/10.17605/OSF.IO/PZ67M

Austerberry, C., Fearon, P., Ronald, A., Leve, L. D., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., & Reiss, D. (2022). Early manifestations of intellectual performance: Evidence that genetic effects on later academic test performance are mediated through verbal performance in early childhood. *Child Development, 93*(2), e188-e206. https://doi.org/10.1111/cdev.13706

Austerberry, C., Mateen, M., Fearon, P., & Ronald, A. (2022). Heritability of psychological traits and developmental milestones in infancy: A systematic review and meta-analysis. *JAMA Network Open, 5*(8), e2227887-e2227887. https://doi.org/10.1001/jamanetworkopen.2022.27887

Austerberry, C., Zayats, T., Ronald, A., Corfield, E. C., Smajlagic, D., Havdahl, A., Andreassen, O. A., Magnus, P., Njølstad, P., & Fearon, P. (2022, February 9). Evocative effects of common genetic variants associated with educational attainment on the early caregiving environment in the Norwegian Mother, Father and Child Cohort Study (MoBa). https://doi.org/10.17605/OSF.IO/84RMN

Bakermans-Kranenburg, M. J., van Uzendoorn, M. H., Bokhorst, C. L., & Schuengel, C. (2004). The importance of shared environment in infant-father attachment: A behavioral genetic study of the attachment q-sort. *The Journal of Family Psychology, 18*(3), 545-549. https://doi.org/10.1037/0893-3200.18.3.545

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

Beaver, K. M., Boutwell, B. B., Barnes, J., Schwartz, J. A., & Connolly, E. J. (2014). A quantitative genetic analysis of the associations among language skills, peer interactions, and behavioral problems in childhood: Results from a sample of twins. *Merrill-Palmer Quarterly, 60*(2), 142-167. https://doi.org/10.13110/merrpalmquar1982.60.2.0142

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75*(2), 81-95. https://doi.org/10.1037/h0025583

Belsky, D. W., Moffitt, T. E., Corcoran, D. L., Domingue, B., Harrington, H., Hogan, S., Houts, R., Ramrakha, S., Sugden, K., Williams, B. S., Poulton, R., & Caspi, A. (2016). The genetics of success: How single-nucleotide polymorphisms associated with educational attainment relate to life-course development. *Psychololgical Science, 27*, 957-972. https://doi.org/10.1177/0956797616643070

Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*(1), 78-117. https://doi.org/10.1177/0049124187016001004

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics, 10*(3), 423-433. https://doi.org/10.1375/twin.10.3.423

Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and psychological measurement, 65*(5), 676-696. https://doi.org/10.1177/0013164404272507

Bird, P. K., McEachan, R. R. C., Mon-Williams, M., Small, N., West, J., Whincup, P., Wright, J., Andrews, E., Barber, S. E., Hill, L. J. B., Lennon, L., Mason, D., Shire, K. A., Waiblinger, D., Waterman, A. H., Lawlor, D. A., & Pickett, K. E. (2019). Growing up in Bradford: Protocol for the age 7–11 follow up of the Born in Bradford birth cohort. *BMC Public Health, 19*(1), 939. https://doi.org/10.1186/s12889-019-7222-2

Bishop, E., Cherny, S. S., Corley, R., Plomin, R., DeFries, J. C., & Hewitt, J. K. (2003). Development genetic analysis of general cognitive ability from 1 to 12 years in a sample of adoptees, biological siblings, and twins. *Intelligence, 31*(1), 31-49. https://doi.org/10.1016/S0160-2896(02)00112-5

Blair, C., Raver, C. C., Berry, D. J., & Family Life Project Investigators. (2014, Feb). Two approaches to estimating the effect of parenting on the development of executive function in early childhood. *Developmental Psychology, 50*(2), 554-565. https://doi.org/10.1037/a0033647

Boivin, M., Pérusse, D., Dionne, G., Saysset, V., Zoccolillo, M., Tarabulsy, G. M., Tremblay, N., & Tremblay, R. E. (2005). The genetic-environmental etiology of parents' perceptions and self-assessed behaviours toward their 5-month-old infants in a large twin and singleton sample. *Journal of Child Psychology and Psychiatry, 46*(6), 612-630. https://doi.org/https://doi.org/10.1111/j.1469-7610.2004.00375.x

Bokhorst, C. L., Bakermans-Kranenburg, M. J., Fearon, R. M., van, I. M. H., Fonagy, P., & Schuengel, C. (2003). The importance of shared environment in mother-infant attachment security: A behavioral genetic study. *Child Development, 74*(6), 1769-1782. https://doi.org/10.1046/j.1467-8624.2003.00637.x

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20*, 115-140. https://doi.org/10.2307/271084

Boomsma, D. I., Orlebeke, J. F., & van Baal, G. C. (1992). The Dutch Twin Register: Growth data on weight and height. *Behavior Genetics, 22*(2), 247-251. https://doi.org/10.1007/BF01067004

Bornstein, M. H. (2014). Human infancy…and the rest of the lifespan. *Annual review of psychology, 65*, 121-158. https://doi.org/10.1146/annurev-psych-120710-100359

Borriello, G. A., Ramos, A. M., Natsuaki, M. N., Reiss, D., Shaw, D. S., Leve, L. D., & Neiderhiser, J. M. (2020). The intergenerational transmission of mathematics achievement in middle childhood: A prospective adoption design. *Developmental Science, e12974*. https://doi.org/10.1111/desc.12974

Bouchard, T. J., Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*(4498), 1055-1059. https://doi.org/10.1126/science.7195071

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

Bouchard, T. J., & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics, 31*(3), 243-273. https://doi.org/10.1023/A:1012294324713

Brant, A. M., Haberstick, B. C., Corley, R. P., Wadsworth, S. J., DeFries, J. C., & Hewitt, J. K. (2009). The developmental etiology of high IQ. *Behavior Genetics, 39*(4), 393-405. https://doi.org/10.1007/s10519-009-9268-x

Brescianini, S., Giampietro, S., Cotichini, R., Lucchini, R., & De Curtis, M. (2012). Genetic and environmental components of neonatal weight gain in preterm infants. *Pediatrics, 129*(2), e455-459. https://doi.org/10.1542/peds.2010-0510

Brescianini, S., Volzone, A., Fagnani, C., Patriarca, V., Grimaldi, V., Lanni, R., Serino, L., Mastroiacovo, P., & Stazi, M. A. (2011). Genetic and environmental factors shape infant sleep patterns: A study of 18-month-old twins. *Pediatrics, 127*(5), e1296-1302. https://doi.org/10.1542/peds.2010-0858

Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science, 24*(9), 1704-1713. https://doi.org/10.1177/0956797613478618

Byrne, R. P., van Rheenen, W., van den Berg, L. H., Veldink, J. H., McLaughlin, R. L., & Project MinE ALS Gwas Consortium. (2020). Dutch population structure across space, time and GWAS design. *Nature Communications, 11*(1), 4556. https://doi.org/10.1038/s41467-020-18418-4

Capron, C., & Duyme, M. (1989). How much can we raise or lower I.Q.? The joint effects genetic and prenatal factors versus postnatal factors. *Journal de Genetic Human, 37*(2), 115-118.

Caramaschi, D., Booij, L., Petitclerc, A., Boivin, M., & Tremblay, R. E. (2012). Genetic and environmental contributions to saliva testosterone levels in male and female infant twins. *Psychoneuroendocrinology, 37*(12), 1954-1959. https://doi.org/10.1016/j.psyneuen.2012.04.008

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research, 1*(2), 245-276.

Cheesman, R., Hunjan, A., Coleman, J. R. I., Ahmadzadeh, Y., Plomin, R., McAdams, T. A., Eley, T. C., & Breen, G. (2020). Comparison of adopted and nonadopted individuals reveals gene–environment interplay for education in the UK Biobank. *Psychological Science, 31*(5), 582-591. https://doi.org/10.1177/0956797620904450

Chen, C. J., Yu, M. W., Wang, C. J., Tong, S. L., Tien, M., Lee, T. Y., Lue, H. C., Huang, F. Y., Lan, C. C., Yang, K. H., & et al. (1990a). Chronological changes in genetic variance and heritability of anthropometric characteristics among Chinese twin infants. *Acta Geneticae Medicae et Gemellologiae, 39*(4), 479-484. https://doi.org/10.1017/s0001566000003706

Chen, C. J., Yu, M. W., Wang, C. J., Tong, S. L., Tien, M., Lee, T. Y., Lue, H. C., Huang, F. Y., Lan, C. C., Yang, K. H., & et al. (1990b). Genetic variance and heritability of temperament among Chinese twin infants. *Acta Geneticae Medicae et Gemellologiae, 39*(4), 485-490. https://doi.org/10.1017/s0001566000003718

Cherny, S., Cardon, L., Fulker, D. W., & DeFries, J. (1992). Differential heritability across levels of cognitive ability. *Behavior Genetics, 22*(2), 153-162. https://doi.org/10.1007/BF01066994

Cherny, S. S., Fulker, D., Emde, R., Robinson, J., Corley, R., Reznick, J., Plomin, R., & DeFries, J. (1994). A developmental-genetic analysis of continuity and change in the Bayley Mental Development Index from 14 to 24 months: The MacArthur Longitudinal Twin Study. *Psychological Science, 5*(6), 354-360. https://doi.org/10.1111/j.1467-9280.1994.tb00285.x

Cherny, S. S., Fulker, D. W., Corley, R. P., Plomin, R., & DeFries, J. C. (1994). Continuity and change in infant shyness from 14 to 20 months. *Behavior Genetics, 24*(4), 365-379. https://doi.org/10.1007/BF01067538

Cheung, M. W. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology, 5*(1521). https://doi.org/https://doi.org/10.3389/fpsyg.2014.01521

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

Cioffi, C. C., Griffin, A. M., Natsuaki, M. N., Shaw, D. S., Reiss, D., Ganiban, J. M., Neiderhiser, J. M., & Leve, L. D. (2021). The role of negative emotionality in the development of child executive function and language abilities from toddlerhood to first grade: An adoption study. *Developmental Psychology, 57*, 347-360. https://doi.org/10.1037/dev0000972

Connolly, E. J., Schwartz, J. A., Nedelec, J. L., Beaver, K. M., & Barnes, J. C. (2015). Different slopes for different folks: Genetic influences on growth in delinquent peer association and delinquency during adolescence. *Journal of Youth and Adolescence, 44*(7), 1413-1427. https://doi.org/10.1007/s10964-015-0299-8

Cooper, R. M., & Zubek, J. P. (1958). Effects of enriched and restricted early environments on the learning ability of bright and dull rats. *Canadian Journal of Psychology, 12*(3), 159-164. https://doi.org/10.1037/h0083747

Corfield, E., Frei, O., Shadrin, A. A., Rahman, Z., Lin, A., Athanasiu, L., Akdeniz, B. C., Hannigan, L., Wootton, R., Austerberry, C., Hughes, A., Tesli, M., Westlye, L. T., Njølstad, P., Magnus, P., Davies, N. M., Appadurai, V., Hemani, G., Hovig, E., Zayats, T., Ask, H., Reichborn-Kjennerud, T., Andreassen, O. A., & Havdahl, A. (2022). The Norwegian Mother, Father, and Child cohort study (MoBa) genotyping data resource: MoBaPsychGen pipeline v.1. Preprint: https://www.biorxiv.org/content/10.1101/2022.06.23.496289v2.

Custodio, R. J., Junior, C. E., Milani, S. L., Simoes, A. L., de Castro, M., & Moreira, A. C. (2007). The emergence of the cortisol circadian rhythm in monozygotic and dizygotic twin infants: the twin-pair synchrony. *Clinical*

Endocrinology, 66(2), 192-197. https://doi.org/10.1111/j.1365-2265.2006.02706.x

D'Onofrio, B. M., Turkheimer, E. N., Eaves, L. J., Corey, L. A., Berg, K., Solaas, M. H., & Emery, R. E. (2003). The role of the Children of Twins design in elucidating causal relations between parent characteristics and child outcomes. *Journal of Child Psychology and Psychiatry, 44*(8), 1130-1144. https://doi.org/https://doi.org/10.1111/1469-7610.00196

Dale, P. S., Dionne, G., Eley, T. C., & Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language, 27*(3), 619-642. https://doi.org/10.1017/s0305000900004281

Darling, J. C., Bamidis, P. D., Burberry, J., & Rudolf, M. C. J. (2020). The first thousand days: Early, integrated and evidence-based approaches to improving child health: Coming to a population near you? *Archives of Disease in Childhood, 105*(9), 837–841. https://doi.org/http://dx.doi.org/10.1136/archdischild-2019-316929

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

Davis, D. W., Finkel, D., Turkheimer, E., & Dickens, W. (2015). Genetic and environmental contributions to behavioral stability and change in children 6-36 months of age using Louisville Twin Study data. *Behavior Genetics, 45*(6), 610-621. https://doi.org/10.1007/s10519-015-9759-x

Davis, O. S., Haworth, C. M., & Plomin, R. (2009a). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science, 20*(10), 1301-1308. https://doi.org/10.1111/j.1467-9280.2009.02433.x

Davis, O. S., Haworth, C. M., & Plomin, R. (2009b). Learning abilities and disabilities: generalist genes in early adolescence. *Cognitive Neuropsychiatry, 14*, 312-331. https://doi.org/10.1080/13546800902797106

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of family psychology, 19*(2), 294.

Davis-Kean, P. E., Tang, S., & Waters, N. E. (2019). Parent education attainment and parenting. In *Handbook of parenting* (pp. 400-420). Routledge.

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79. https://doi.org/10.1177/1529100610387081

Derks, E. M., Dolan, C. V., & Boomsma, D. I. (2006). A test of the equal environment assumption (EEA) in multivariate twin studies. *Twin Research and Human Genetics, 9*(3), 403-411. https://doi.org/10.1375/twin.9.3.403

Dick, D. M. (2011). Gene-environment interaction in psychological traits and disorders. *Annual Review of Clinical Psychology, 7*(1), 383-409. https://doi.org/10.1146/annurev-clinpsy-032210-104518

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346. https://doi.org/10.1037/0033-295X.108.2.346

DiLalla, L. F., & Bishop, E. G. (1996). Differential maternal treatment of infant twins: Effects of infant behaviors. *Behavior Genetics, 26*(6), 535-542. https://doi.org/10.1007/BF02361226

DiLalla, L. F., Kagan, J., & Reznick, J. S. (1994). Genetic etiology of behavioral inhibition among 2-year-old children. *Infant Behavior and Development, 17*, 405-412. https://doi.org/10.1016/0163-6383(94)90032-9

Dionne, G., Dale, P. S., Boivin, M., & Plomin, R. (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development, 74*(2), 394-412. https://doi.org/1467-8624.7402005

Dionne, G., Touchette, E., Forget-Dubois, N., Petit, D., Tremblay, R. E., Montplaisir, J. Y., & Boivin, M. (2011). Associations between sleep-wake consolidation and language development in early childhood: A longitudinal twin study. *Sleep, 34*(8), 987-995. https://doi.org/10.5665/sleep.1148

Dionne, G., Tremblay, R., Boivin, M., Laplante, D., & Perusse, D. (2003). Physical aggression and expressive vocabulary in 19-month-old twins. *Developmental Psychology, 39*(2), 261-273. https://doi.org/10.1037//0012-1649.39.2.261

Domingue, B. W., & Fletcher, J. (2020). Separating measured genetic and environmental effects: Evidence linking parental genotype and adopted child outcomes. *Behavior Genetics, 50*, 301-309. https://doi.org/10.1007/s10519-020-10000-4

Dubois, L., Girard, M., Girard, A., Tremblay, R., Boivin, M., & Perusse, D. (2007). Genetic and environmental influences on body size in early childhood: A twin birth-cohort study. *Twin Research and Human Genetics, 10*(3), 479-485. https://doi.org/10.1375/twin.10.3.479

Dubois, L., Ohm Kyvik, K., Girard, M., Tatone-Tokuda, F., Perusse, D., Hjelmborg, J., Skytthe, A., Rasmussen, F., Wright, M. J., Lichtenstein, P., & Martin, N. G. (2012). Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: An international study of over 12,000 twin pairs. *PLoS One, 7*(2), e30153. https://doi.org/10.1371/journal.pone.0030153

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428-1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of “dull children”?: A late adoption study. *Proceedings of the National Academy of Sciences, 96*(15), 8790-8794. https://doi.org/10.1073/pnas.96.15.8790

Eaves, L. J., & Rao, D. C. (1984). The resolution of genotype × environment interaction in segregation analysis of nuclear families. *Genetic Epidemiology, 1*(3), 215-228. https://doi.org/10.1002/gepi.1370010302

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/https://doi.org/10.1136/bmj.315.7109.629

Elam, K. K., Harold, G. T., Neiderhiser, J. M., Reiss, D., Shaw, D. S., Natsuaki, M. N., Gaysina, D., Barrett, D., & Leve, L. D. (2014). Adoptive parent hostility and children's peer behavior problems: Examining the role of genetically informed child attributes on adoptive parent behavior. *Developmental Psychology, 50*(5), 1543-1552. https://doi.org/10.1037/a0035470

Emde, R. N., Plomin, R., Robinson, J. A., Corley, R., DeFries, J., Fulker, D. W., Reznick, J. S., Campos, J., Kagan, J., & Zahn-Waxler, C. (1992). Temperament, emotion, and cognition at fourteen months: The MacArthur Longitudinal Twin Study. *Child Development, 63*(6), 1437-1455. https://doi.org/10.2307/1131567

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 8*(3), 430-457. https://doi.org/10.1207/S15328007SEM0803_5

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Psychometric properties of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies, 15*(5), 595-614. https://doi.org/10.1007/s10826-006-9036-y

Evans, D. M., & Davey Smith, G. (2015). Mendelian randomization: New applications in the coming age of hypothesis-free causality. *Annual Review of Genomics and Human Genetics, 16*, 327-350. https://doi.org/10.1146/annurev-genom-090314-050016

Falck-Ytter, T., Hamrefors, L., Siqueiros Sanches, M., Portugal, A. M., Taylor, M., Li, D., Viktorsson, C., Hardiansyah, I., Myers, L., Westberg, L., Bölte, S., Tammimies, K., & Ronald, A. (2021). The Babytwins Study Sweden (BATSS): A multi-method infant twin study of genetic and environmental factors influencing infant brain and behavioral development. *Twin Research and Human Genetics, 24*(4), 217-227. https://doi.org/10.1017/thg.2021.34

Falconer, D. S. (1960). *Introduction to quantitative genetics*. Ronald Press Co.

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1-22. https://doi.org/10.1023/A:1009048817385

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

Fearon, P., Reiss, D., Leve, L. D., Shaw, D. S., Scaramella, L. V., Ganiban, J. M., & Neiderhiser, J. M. (2015). Child-evoked maternal negativity from 9 to 27 months: Evidence of gene–environment correlation and its moderation by marital distress. *Development and Psychopathology, 27*(4pt1), 1251-1265. https://doi.org/10.1017/S0954579414000868

Finkel, D., & Matheny, A. P., Jr. (2000). Genetic and environmental influences on a measure of infant attachment security. *Twin Research, 3*(4), 242-250. https://doi.org/10.1375/136905200320565210

Fisher, A., van Jaarsveld, C. H., Llewellyn, C. H., & Wardle, J. (2012). Genetic and environmental influences on infant sleep. *Pediatrics, 129*(6), 1091-1096. https://doi.org/10.1542/peds.2011-1571

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh, 52*(2), 399-433.

Flom, M., & Saudino, K. J. (2017). Callous-unemotional behaviors in early childhood: Genetic and environmental contributions to stability and change. *Development and Psychopathology, 29*(4), 1227-1234. https://doi.org/10.1017/s0954579416001267

Flom, M., & Saudino, K. J. (2018). Do Genetic Factors Explain the Links Between Callous-Unemotional, Attention Hyperactivity and Oppositional Defiant Problems in Toddlers? *Journal of Abnormal Child Psychology, 46*(6), 1217-1228. https://doi.org/10.1007/s10802-017-0361-y

Flom, M., Wang, M., Uccello, K. J., & Saudino, K. J. (2018). Parent- and observer-rated positive affect in early childhood: Genetic overlap and environmental specificity. *Behavior Genetics, 48*(6), 432-439. https://doi.org/10.1007/s10519-018-9924-0

Flom, M., White, D., Ganiban, J., & Saudino, K. J. (2019). Longitudinal links between callous-unemotional behaviors and parenting in early childhood: A genetically informed design. *The Journal of the American Academy of Child and Adolescent Psychiatry, 59*(3), 401-409. https://doi.org/10.1016/j.jaac.2019.03.013

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*(2), 171. https://doi.org/10.1037/0033-2909.101.2.171

Forget-Dubois, N., Boivin, M., Dionne, G., Pierce, T., Tremblay, R. E., & Perusse, D. (2007). A longitudinal twin study of the genetic and environmental etiology of maternal hostile-reactive behavior during infancy and toddlerhood. *Infant Behavior & Development, 30*(3), 453-465. https://doi.org/10.1016/j.infbeh.2006.12.005

Frick, P. J. (1991). *Alabama Parenting Questionnaire*. University of Alabama.

Frick, P. J., Christian, R. E., & Wootton, J. M. (1999). Age trends in the association between parenting practices and conduct problems. *Behavior Modification, 23*(1), 106-128. https://doi.org/10.1177/0145445599231005

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: A behavioral genetic analysis. *Developmental Psychology, 47*(5), 1410-1430. https://doi.org/10.1037/a0023750

Fritz, M. S., Kenny, D. A., & MacKinnon, D. P. (2016). The combined effects of measurement error and omitting confounders in the single-mediator model. *Multivariate behavioral research, 51*(5), 681-697. https://doi.org/10.1080/00273171.2016.1224154

Fujisawa, K. K., Ozaki, K., Suzuki, K., Yamagata, S., Kawahashi, I., & Ando, J. (2012). Genetic and environmental relationships between head circumference growth in the first year of life and sociocognitive development in the second year: A longitudinal twin study. *Developmental Science, 15*(1), 99-112. https://doi.org/10.1111/j.1467-7687.2011.01097.x

Gagne, J. R., Asherson, P., & Saudino, K. J. (2020). A twin study of inhibitory control at age two and ADHD behavior problems at age three. *Behavior Genetics, 50*(4), 289-300. https://doi.org/10.1007/s10519-020-09997-5

Gagne, J. R., & Goldsmith, H. H. (2011). A longitudinal analysis of anger and inhibitory control in twins from 12 to 36 months of age. *Developmental Science, 14*(1), 112-124. https://doi.org/10.1111/j.1467-7687.2010.00969.x

Gagne, J. R., & Saudino, K. J. (2010). Wait for it! A twin study of inhibitory control in early childhood. *Behavior Genetics, 40*(3), 327-337. https://doi.org/10.1007/s10519-009-9316-6

Gagne, J. R., & Saudino, K. J. (2016). The development of inhibitory control in early childhood: A twin study from 2-3 years. *Developmental Psychology, 52*(3), 391-399. https://doi.org/10.1037/dev0000090

Gagne, J. R., Saudino, K. J., & Asherson, P. (2011). The genetic etiology of inhibitory control and behavior problems at 24 months of age. *Journal of Child Psychology and Psychiatry, 52*(11), 1155-1163. https://doi.org/10.1111/j.1469-7610.2011.02420.x

Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000a). Sex differences in early verbal and non-verbal cognitive development.

Developmental Science, 3(2), 206-215. https://doi.org/10.1111/1467-7687.00114

Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000b). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science, 3*, 206-215. https://doi.org/10.1111/1467-7687.00114

Gao, W., Cao, W., Lv, J., Yu, C., Wu, T., Wang, S., Meng, L., Wang, D., Wang, Z., Pang, Z., Yu, M., Wang, H., Wu, X., Dong, Z., Wu, F., Jiang, G., Wang, X., Liu, Y., Deng, J., Lu, L., & Li, L. (2019). The Chinese National Twin Registry: A 'gold mine' for scientific research. *Journal of Internal Medicine, 286*(3), 299-308. https://doi.org/https://doi.org/10.1111/joim.12926

Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications, 10*(1), 1776. https://doi.org/10.1038/s41467-019-09718-5

Ge, X., Conger, R. D., Cadoret, R. J., Neiderhiser, J. M., Yates, W., Troughton, E., & Stewart, M. A. (1996). The developmental interface between nature and nurture: A mutual influence model of child antisocial behavior and parent behaviors. *Developmental Psychology, 32*(4), 574-589. https://doi.org/10.1037/0012-1649.32.4.574

Ge, X., Natsuaki, M. N., Martin, D. M., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Villareal, G., Scaramella, L., Reid, J. B., & Reiss, D. (2008). Bridging the divide: Openness in adoption and postadoption psychosocial adjustment among birth and adoptive parents. *The Journal of Family Psychology, 22*, 529-540. https://doi.org/10.1037/a0012817

German, A., Livshits, G., Peter, I., Malkin, I., Dubnov, J., Akons, H., Shmoish, M., & Hochberg, Z. (2015). Environmental rather than genetic factors determine the variation in the age of the infancy to childhood transition: A twins study. *The Journal of Pediatrics, 166*(3), 731-735. https://doi.org/10.1016/j.jpeds.2014.11.047

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*, 129-153. https://doi.org/10.1016/0010-0277(94)90068-x.

Gilmore, J. H., Schmitt, J. E., Knickmeyer, R. C., Smith, J. K., Lin, W., Styner, M., Gerig, G., & Neale, M. C. (2010). Genetic and environmental contributions to neonatal brain structure: A twin study. *Human Brain Mapping, 31*(8), 1174-1182. https://doi.org/10.1002/hbm.20926

Goetghebuer, T., Ota, M. O., Kebbeh, B., John, M., Jackson-Sillah, D., Vekemans, J., Marchant, A., Newport, M., & Weiss, H. A. (2003). Delay in motor development of twins in Africa: A prospective cohort study. *Twin Research, 6*(4), 279-284. https://doi.org/10.1375/136905203322296629

Goldsmith, H. H., & Gottesman, I. I. (1981). Origins of variation in behavioral style: A longitudinal study of temperament in young twins. *Child Development, 52*(1), 91-103. https://doi.org/10.2307/1129218

Goldsmith, H. H., Lemery, K. S., Buss, K. A., & Campos, J. J. (1999). Genetic analyses of focal aspects of infant temperament. *Developmental Psychology, 35*(4), 972-985.

Goldsmith, H. H., & Rothbart, M. K. (1999). *The laboratory temperament assessment battery (LAB-TAB): Prelocomotor version, technical manual* (3.1 ed.). Department of Psychology, University of Wisconsin.

Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Institute for the Development of Educational Achievement.

Hadley, P. A., & Rice, M. L. (1993). Parental judgments of preschoolers' speech and language development: A resource for assessment and IEP planning. *Topics in Speech and Language, 14*(04), 278-288. https://doi.org/10.1055/s-2008-1064177

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015, Mar). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102-116. https://doi.org/10.1037/a0038889

Harold, G. T., Leve, L. D., Barrett, D., Elam, K. K., Neiderhiser, J. M., Natsuaki, M. N., Shaw, D. S., Reiss, D., & Thapar, A. (2013). Biological and rearing mother influences on child ADHD symptoms: Revisiting the developmental interface between nature and nurture. *Journal of Child Psychology and Psychiatry, 54*(10), 1038-1046. https://doi.org/10.1111/jcpp.12100

Hawks, Z. W., Marrus, N., Glowinski, A. L., & Constantino, J. N. (2018, Mar 16). Early origins of autism comorbidity: Neuropsychiatric traits correlated in childhood are independent in infancy. *Journal of Abnormal Child Psychology, 47*(2), 369-379. https://doi.org/10.1007/s10802-018-0410-1

Haworth, C. M., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J., van Beijsterveldt, C. E., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S., Kovas, Y., Corley, R. P., Defries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010, Nov). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*(11), 1112–1120. https://doi.org/https://doi.org/10.1038/mp.2009.55

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S. P., Kovas, Y., Corley, R. P., DeFries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, S. A., Petrill, S. A., Lubinski, D., & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*(11), 1112-1120. https://doi.org/10.1038/mp.2009.55

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiry, 46*(3), 289. https://doi.org/https://doi.org/10.1111/j.1465-7295.2008.00163.x

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Herle, M., Fildes, A., & Llewellyn, C. (2018, Oct). Emotional eating is learned not inherited in children, regardless of obesity risk [Empirical Study; Experimental Replication; Mathematical Model; Quantitative Study; Twin Study]. *Pediatric Obesity, 13*(10), 628-631. https://doi.org/http://dx.doi.org/10.1111/ijpo.12428

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560. https://doi.org/https://doi.org/10.1136/bmj.327.7414.557

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

Horn, J. M. (1983). The Texas Adoption Project: Adopted children and their intellectual resemblance to biological and adoptive parents. *Child Development, 54*(2), 268-275. https://doi.org/10.2307/1129690

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hummer, R. A., & Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States. *Population bulletin, 68*(1), 1-16.

Hur, Y. M. (2005). Genetic and environmental influences on birthweight in a sample of Korean twins. *Journal of Korean Medical Science, 20*(3), 355-360. https://doi.org/10.3346/jkms.2005.20.3.355

Hur, Y. M., Jeong, H. U., Kang, M. C., Ajose, F., Kim, J. W., Beck, J. J., Hottenga, J. J., Mbarek, H., Finnicum, C. T., Ehli, E. A., Martin, N. G., de Geus, E. J., Boomsma, D. I., Davies, G. E., & Bates, T. (2019). The Nigerian Twin and Sibling Registry: An update. *Twin Research and Human Genetics, 22*(6), 637-640. https://doi.org/10.1017/thg.2019.110

Hur, Y. M., Kang, M. C., Jeong, H.-U., Kang, I. C., & Kim, J. W. (2019). The South Korean Twin Registry. *Twin Research and Human Genetics, 22*(6), 606-608. https://doi.org/10.1017/thg.2019.115

Hur, Y. M., Luciano, M., Martin, N. G., Boomsma, D. I., Iacono, W. G., McGue, M., Shin, J. S., Jun, J. K., Ooki, S., van Beijsterveldt, C. E., & Han, J. Y. (2005). A comparison of twin birthweight data from Australia, the Netherlands, the United States, Japan, and South Korea: Are genetic and environmental variations in birthweight similar in Caucasians and East Asians? *Twin Research and Human Genetics, 8*(6), 638-648. https://doi.org/10.1375/183242705774860169

Ilott, N., Saudino, K. J., Wood, A., & Asherson, P. (2010). A genetic study of ADHD and activity level in infancy. *Genes Brain and Behavior, 9*(3), 296-304. https://doi.org/10.1111/j.1601-183X.2009.00560.x

Ilott, N. E., Saudino, K. J., & Asherson, P. (2010). Genetic influences on attention deficit hyperactivity disorder symptoms from age 2 to 3: A quantitative and molecular genetic investigation. *BMC Psychiatry, 10*, 102. https://doi.org/10.1186/1471-244x-10-102

Isungset, M. A., Conley, D. C., Zachrisson, H. D., Ystrøm, E., Havdahl, A., Njølstad, P. R., & Lyngstad, T. H. (2021). *Social and genetic effects on educational performance in early adolescence*.

Jackson, D. B. (2016). The association between breastfeeding duration and attachment: A genetically informed analysis. *Breastfeeding Medicine, 11*(6), 297-304. https://doi.org/10.1089/bfm.2016.0036

Jha, S. C., Xia, K., Schmitt, J. E., Ahn, M., Girault, J. B., Murphy, V. A., Li, G., Wang, L., Shen, D., Zou, F., Zhu, H., Styner, M., Knickmeyer, R. C., & Gilmore, J. H. (2018). Genetic influences on neonatal cortical thickness and surface area. *Human Brain Mapping, 39*(12), 4998-5013. https://doi.org/10.1002/hbm.24340

Johnson, A. D., Martin, A., Brooks-Gunn, J., & Petrill, S. A. (2008). Order in the House! Associations among Household Chaos, the Home Literacy Environment, Maternal Reading Ability, and Children's Early Reading. *Merrill Palmer Q (Wayne State Univ Press), 54*(4), 445-472. https://doi.org/10.1353/mpq.0.0009

Johnson, L., Llewellyn, C. H., van Jaarsveld, C. H., Cole, T. J., & Wardle, J. (2011). Genetic and environmental influences on infant growth: Prospective analysis of the Gemini twin birth cohort. *PLoS One, 6*(5), e19918. https://doi.org/10.1371/journal.pone.0019918

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5*(5), 602-619. https://doi.org/10.1177/0193841X8100500502

Kalil, A., Ryan, R., & Corey, M. (2012). Diverging destinies: Maternal education and the developmental gradient in time with children. *Demography, 49*(4), 1361-1383. https://doi.org/10.1007/s13524-012-0129-5

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215-227. https://doi.org/10.1080/02796015.1996.12085812

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. John Wiley & Sons.

Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: A systematic review. *Psychological Medicine, 37*(5), 615-626. https://doi.org/10.1017/S0033291706009524

Kendler, K. S., & Neale, M. C. (2010). Endophenotype: A conceptual analysis. *Molecular Psychiatry, 15*(8), 789-797. https://doi.org/10.1038/mp.2010.8

Kendler, K. S., Turkheimer, E., Ohlsson, H., Sundquist, J., & Sundquist, K. (2015). Family environment and the malleability of cognitive ability: A Swedish national home-reared and adopted-away cosibling control study. *Proceedings of the National Academy of Sciences, 112*(15), 4612-4617. https://doi.org/10.1073/pnas.1417106112

Klahr, A. M., Burt, S. A., Leve, L. D., Shaw, D. S., Ganiban, J. M., Reiss, D., & Neiderhiser, J. M. (2017). Birth and adoptive parent antisocial behavior and parenting: A study of evocative gene–environment correlation. *Child Development, 88*(2), 505-513. https://doi.org/10.1111/cdev.12619

Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? *Journal of Marriage and Family, 56*(2), 441-455. https://doi.org/10.2307/353111

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research, 41*, 627-641. https://doi.org/10.1044/jslhr.4103.627

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. The Guilford Press.

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*.

Knafo, A., & Plomin, R. (2006a). Parental discipline and affection and children's prosocial behavior: Genetic and environmental links. *Journal of Personality and Social Psychology, 90*(1), 147. https://doi.org/10.1037/0022-3514.90.1.147

Knafo, A., & Plomin, R. (2006b). Prosocial behavior from early to middle childhood: Genetic and environmental influences on stability and change. *Developmental Psychology, 42*(5), 771-786. https://doi.org/10.1037/0012-1649.42.5.771

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*, 490-507.

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology, 36*, 220-232.

Koellinger, P. D., & Harden, K. P. (2018). Using nature to understand nurture. *Science, 359*(6374), 386-387. https://doi.org/10.1126/science.aar6429

Koeppen-Schomerus, G., Spinath, F. M., & Plomin, R. (2003). Twins and non-twin siblings: Different estimates of shared environmental influence in early childhood. *Twin Research, 6*(2), 97-105. https://doi.org/10.1375/136905203321536227

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

Koran, J. (2020). Indicators per factor in confirmatory factor analysis: More is not always better. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(5), 765-772. https://doi.org/10.1080/10705511.2019.1706527

Kosik, R., Mandell, G., Fan, A. P., Nguyen, T., Chen, J., & Eaton, W. (2018). The association between childhood educational attainment and adult mental health and status: A thirty-year longitudinal follow up study. *The European Journal of Psychiatry, 32*(2), 53-62. https://doi.org/10.1016/j.ejpsy.2018.01.001

Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*(10), 2048-2056. https://doi.org/10.1177/0956797613486982

Krapohl, E., Hannigan, L. J., Pingault, J.-B., Patel, H., Kadeva, N., Curtis, C., Breen, G., Newhouse, S. J., Eley, T. C., O'Reilly, P. F., & Plomin, R. (2017). Widespread covariation of early environmental exposures and trait-associated polygenic variation. *Proceedings of the National Academy of Sciences, 114*(44), 11727-11732. https://doi.org/10.1073/pnas.1707178114

Kumsta, R., Kreppner, J., Kennedy, M., Knights, N., Rutter, M., & Sonuga-Barke, E. (2015). Psychological consequences of early global deprivation: An overview of findings from the English & Romanian Adoptees study. *European Psychologist, 20*(2), 138. https://doi.org/10.1027/1016-9040/a000227

Kuntsi, J., Rijsdijk, F., Ronald, A., Asherson, P., & Plomin, R. (2005). Genetic influences on the stability of attention-deficit/hyperactivity disorder symptoms from early to middle childhood. *Biological Psychiatry, 57*(6), 647-654. https://doi.org/10.1016/j.biopsych.2004.12.032

Lacourse, E., Boivin, M., Brendgen, M., Petitclerc, A., Girard, A., Vitaro, F., Paquin, S., Ouellet-Morin, I., Dionne, G., & Tremblay, R. E. (2014). A longitudinal twin study of physical aggression during early childhood: Evidence for a developmentally dynamic genome. *Psychological Medicine, 44*(12), 2617-2627. https://doi.org/10.1017/s0033291713003218

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., Yengo, L., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Hinds, D. A., McCreight, J. C., Huber, K. E., Litterman, N. K., McIntyre, M. H., Mountain, J. L., Noblin, E. S., Northover, C. A. M., Pitts, S. J., Sathirapongsasuti, J. F., Sazonova, O. V., Shelton, J. F., Shringarpure, S., Tian, C., Vacic, V., Wilson, C. H., Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., Turley, P., Chen, G.-B., Emilsson, V., Meddens, S. F. W., Oskarsson, S., Pickrell, J. K., Thom, K., Timshel, P., Vlaming, R., Abdellaoui, A., Ahluwalia, T. S., Bacelis, J., Baumbach, C., Bjornsdottir, G., Brandsma, J. H., Concas, M. P., Derringer, J., Furlotte, N. A., Galesloot, T. E., Girotto, G., Gupta, R., Hall, L. M., Harris, S. E., Hofer, E., Horikoshi, M., Huffman, J. E., Kaasik, K., Kalafati, I. P., Karlsson, R., Kong, A., Lahti, J., van der Lee, S. J., Leeuw, C., Lind, P.

A., Lindgren, K.-O., Liu, T., Mangino, M., Marten, J., Mihailov, E., Miller, M. B., van der Most, P. J., Oldmeadow, C., Payton, A., Pervjakova, N., Peyrot, W. J., Qian, Y., Raitakari, O., Rueedi, R., Salvi, E., Schmidt, B., Schraut, K. E., Shi, J., Smith, A. V., Poot, R. A., St Pourcain, B., Teumer, A., Thorleifsson, G., Verweij, N., Vuckovic, D., Wellmann, J., Westra, H.-J., Yang, J., Zhao, W., Zhu, Z., Alizadeh, B. Z., Amin, N., Bakshi, A., Baumeister, S. E., Biino, G., Bønnelykke, K., Boyle, P. A., Campbell, H., Cappuccio, F. P., Davies, G., De Neve, J.-E., Deloukas, P., Demuth, I., Ding, J., Eibich, P., Eisele, L., Eklund, N., Evans, D. M., Faul, J. D., Feitosa, M. F., Forstner, A. J., Gandin, I., Gunnarsson, B., Halldórsson, B. V., Harris, T. B., Heath, A. C., Hocking, L. J., Holliday, E. G., Homuth, G., Horan, M. A., Hottenga, J.-J., de Jager, P. L., Joshi, P. K., Jugessur, A., Kaakinen, M. A., Kähönen, M., Kanoni, S., Keltigangas-Järvinen, L., Kiemeney, L. A. L. M., Kolcic, I., Koskinen, S., Kraja, A. T., Kroh, M., Kutalik, Z., Latvala, A., Launer, L. J., Lebreton, M. P., Levinson, D. F., Lichtenstein, P., Lichtner, P., Liewald, D. C. M., Life Lines Cohort Study, Loukola, A., Madden, P. A., Mägi, R., Mäki-Opas, T., Marioni, R. E., Marques-Vidal, P., Meddens, G. A., McMahon, G., Meisinger, C., Meitinger, T., Milaneschi, Y., Milani, L., Montgomery, G. W., Myhre, R., Nelson, C. P., Nyholt, D. R., Ollier, W. E. R., Palotie, A., Paternoster, L., Pedersen, N. L., Petrovic, K. E., Porteous, D. J., Räikkönen, K., Ring, S. M., Robino, A., Rostapshova, O., Rudan, I., Rustichini, A., Salomaa, V., Sanders, A. R., Sarin, A.-P., Schmidt, H., Scott, R. J., Smith, B. H., Smith, J. A., Staessen, J. A., Steinhagen-Thiessen, E., Strauch, K., Terracciano, A., Tobin, M. D., Ulivi, S., Vaccargiu, S., Quaye, L., van Rooij, F. J. A., Venturini, C., Vinkhuyzen, A. A. E., Völker, U., Völzke, H., Vonk, J. M., Vozzi, D., Waage, J., Ware, E. B., Willemsen, G., Attia, J. R., Bennett, D. A., Berger, K., Bertram, L., Bisgaard, H., Boomsma, D. I., Borecki, I. B., Bültmann, U., Chabris, C. F., Cucca, F., Cusi, D., Deary, I. J., Dedoussis, G. V., van Duijn, C. M., Eriksson, J. G., Franke, B., Franke, L., Gasparini, P., Gejman, P. V., Gieger, C., Grabe, H.-J., Gratten, J., Groenen, P. J. F., Gudnason, V., van der Harst, P., Hayward, C., Hinds, D. A., Hoffmann, W., Hyppönen, E., Iacono, W. G., Jacobsson, B., Järvelin, M.-R., Jöckel, K.-H., Kaprio, J., Kardia, S. L. R., Lehtimäki, T., Lehrer, S. F., Magnusson, P. K. E., Martin, N. G., McGue, M., Metspalu, A., Pendleton, N., Penninx, B. W. J. H., Perola, M., Pirastu, N., Pirastu, M., Polasek, O., Posthuma, D., Power, C., Province, M. A., Samani, N. J., Schlessinger, D., Schmidt, R., Sørensen, T. I. A., Spector, T. D., Stefansson, K., Thorsteinsdottir, U., Thurik, A. R., Timpson, N. J., Tiemeier, H., Tung, J. Y., Uitterlinden, A. G., Vitart, V., Vollenweider, P., Weir, D. R., Wilson, J. F., Wright, A. F., Conley, D. C., Krueger, R. F., Smith, G. D., Hofman, A., Laibson, D. I., Medland, S. E., Meyer, M. N., Yang, J., Johannesson, M., Visscher, P. M., Esko, T., Koellinger, P. D., Cesarini, D., 23andMe Research Team, Cogent, & Social Science Genetic Association Consortium. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics, 50*(8), 1112-1121. https://doi.org/10.1038/s41588-018-0147-3

Lejarraga, H. (2012). Growth in infancy and childhood: A pediatric approach. In N. Cameron & B. Bogin (Eds.), *Human Growth and Development* (Second ed., pp. 23–56). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-383882-7.00002-7

Lettre, G. (2011). Recent progress in the study of the genetics of height. *Human Genetics, 129*(5), 465-472. https://doi.org/10.1007/s00439-011-0969-x

Leve, L. D., DeGarmo, D. S., Bridgett, D. J., Neiderhiser, J. M., Shaw, D. S., Harold, G. T., Natsuaki, M. N., & Reiss, D. (2013). Using an adoption design to separate genetic, prenatal, and temperament influences on toddler executive function. *Developmental Psychology, 49*, 1045-1057. https://doi.org/10.1037/a0029390

Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics, 22*(6), 716-727. https://doi.org/10.1017/thg.2019.66

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J., Natsuaki, M. N., & Reiss, D. (2013a, Feb). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*(1), 412–423. https://doi.org/https://doi.org/10.1017/thg.2012.126

Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J. M., Natsuaki, M. N., & Reiss, D. (2013b). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics, 16*(1), 412-423. https://doi.org/10.1017/thg.2012.126

Levine, R. S., Hennekens, C. H., & Jesse, M. J. (1987). Genetic variance of weight and length in infant twins. *American Journal of Epidemiology, 126*(5), 929-935. https://doi.org/10.1093/oxfordjournals.aje.a114730

Liu, Q., Yu, C., Gao, W., Cao, W., Lyu, J., Wang, S., Pang, Z., Cong, L., Dong, Z., Wu, F., Wang, H., Wu, X., Jiang, G., Wang, B., & Li, L. (2015). Genetic and environmental effects on weight, height, and BMI under 18 years in a Chinese population-based twin sample. *Twin Research and Human Genetics, 18*(5), 571-580. https://doi.org/10.1017/thg.2015.63

Livshits, G., Peter, I., Vainder, M., & Hauspie, R. (2000). Genetic analysis of growth curve parameters of body weight, height and head circumference. *Annals of Human Biology, 27*(3), 299-312. https://doi.org/10.1080/030144600282181

Llewellyn, C. H., van Jaarsveld, C. H., Johnson, L., Carnell, S., & Wardle, J. (2010). Nature and nurture in infant appetite: Analysis of the Gemini twin birth cohort. *The American Journal of Clinical Nutrition, 91*(5), 1172-1179. https://doi.org/10.3945/ajcn.2009.28868

Lonigan, C. J., Allan, N. P., & Lerner, M. D. (2011). Assessment of preschool early literacy skills: Linking children's educational needs with empirically supported instructional activities. *Psychology in the Schools, 48*, 488-501. https://doi.org/10.1002/pits.20569

Lonigan, C. J., Wagner, R. K., & Torgesen, J. K. (2007). *Test of preschool early literacy examiner's manual*. ProED.

Lugo-Gil, J., & Tamis-LeMonda, C. S. (2008). Family resources and parenting quality: Links to children's cognitive development across the first 3 years. *Child Development, 79*(4), 1065-1085. https://doi.org/10.1111/j.1467-8624.2008.01176.x

MacGillivray, I., Campbell, D. M., & Thompson, B. (1988). *Twinning and twins*. Wiley.

Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., & Jenkins, J. M. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics, 144*(4). https://doi.org/10.1542/peds.2018-3556

Magnus, P., Birke, C., Vejrup, K., Haugan, A., Alsaker, E., Daltveit, A. K., Handal, M., Haugen, M., Høiseth, G., Knudsen, G. P., Paltiel, L., Schreuder, P., Tambs, K., Vold, L., & Stoltenberg, C. (2016). Cohort profile update: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology, 45*(2), 382-388. https://doi.org/10.1093/ije/dyw029

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjaerven, R., & Stoltenberg, C. (2006). Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology, 35*(5), 1146-1150. https://doi.org/10.1093/ije/dyl170

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

Marceau, K., De Araujo-Greecher, M., Miller, E. S., Massey, S. H., Mayes, L. C., Ganiban, J. M., Reiss, D., Shaw, D. S., Leve, L. D., & Neiderhiser, J. M. (2016). The perinatal risk index: Early risks experienced by domestic adoptees in the United States. *PLoS One, 11*, e0150486. https://doi.org/10.1371/journal.pone.0150486

Marceau, K., Horwitz, B. N., Narusyte, J., Ganiban, J. M., Spotts, E. L., Reiss, D., & Neiderhiser, J. M. (2013). Gene–environment correlation underlying the association between parental negativity and adolescent externalizing problems. *Child Development, 84*(6), 2031-2046. https://doi.org/10.1111/cdev.12094

Marrus, N., Glowinski, A. L., Jacob, T., Klin, A., Jones, W., Drain, C. E., Holzhauer, K. E., Hariprasad, V., Fitzgerald, R. T., Mortenson, E. L., Sant, S. M., Cole, L., Siegel, S. A., Zhang, Y., Agrawal, A., Heath, A. C., & Constantino, J. N. (2015). Rapid video-referenced ratings of reciprocal social behavior in toddlers: A twin study. *Journal of Child Psychology and Psychiatry, 56*(12), 1338-1346. https://doi.org/10.1111/jcpp.12391

Marrus, N., Grant, J. D., Harris-Olenak, B., Albright, J., Bolster, D., Haber, J. R., Jacob, T., Zhang, Y., Heath, A. C., Agrawal, A., Constantino, J. N., Elison, J. T., & Glowinski, A. L. (2020). Genetic architecture of reciprocal social behavior in toddlers: Implications for heterogeneity in the early origins of autism spectrum disorder. *Development and Psychopathology, 32*(4), 1190-1205. https://doi.org/10.1017/s0954579420000723

Marrus, N., Kennon-McGill, S., Harris, B., Zhang, Y., Glowinski, A. L., & Constantino, J. N. (2018, Mar 14). Use of a video scoring anchor for rapid serial assessment of social communication in toddlers. *Journal of Visualized Experiments*(133). https://doi.org/10.3791/57041

Matheny, A. P. (1983). A longitudinal twin study of stability of components from Bayley's Infant Behavior Record. *Child Development, 54*, 356-360. https://doi.org/10.2307/1129696

Matheny, A. P., Jr. (1980). Bayley's infant behavior record: Behavioral components and twin analyses. *Child Development, 51*(4), 1157-1167. https://doi.org/10.2307/1129557

Matheny, A. P., Jr. (1984). Twin similarity in the developmental transformations of infant temperament as measured in a multi-method, longitudinal study. *Acta Geneticae Medicae et Gemellologiae, 33*(2), 181-189. https://doi.org/10.1017/s0001566000007212

Matheny, A. P., Jr. (1989). Children's behavioral inhibition over age and across situations: Genetic similarity for a trait during change. *Journal of Personality, 57*(2), 215-235. https://doi.org/10.1111/j.1467-6494.1989.tb00481.x

Matheny, A. P., Jr., Dolan, A. B., & Wilson, R. S. (1976). Twins: Within-pair similarity on Bayley's Infant Behavior Record. *The Journal of Genetic Psychology, 128*(2d Half), 263-270. https://doi.org/10.1080/00221325.1976.10533996

Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of Applied Developmental Psychology, 16*(3), 429-444. https://doi.org/10.1016/0193-3973(95)90028-4

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age

25 educational outcomes. *Early Childhood Research Quarterly, 28*, 314-324. https://doi.org/10.1016/j.ecresq.2012.07.008

McGue, M., Bouchard Jr., T. J., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (Vol. 1, pp. 59-76). American Psychological Association.

McGue, M., Keyes, M., Sharma, A., Elkins, I., Legrand, L., Johnson, W., & Iacono, W. G. (2007). The environments of adopted and non-adopted youth: Evidence on range restriction from the Sibling Interaction and Behavior Study (SIBS). *Behavior Genetics, 37*(3), 449-462. https://doi.org/10.1007/s10519-007-9142-7

McNeil, T. F., Cantor-Graae, E., & Sjostrom, K. (1994). Obstetric complications as antecedents of schizophrenia: Empirical effects of using different obstetric complication scales. *Journal of Psychiatric Research, 28*, 519-530. https://doi.org/10.1016/0022-3956(94)90042-6

McVean, G. A., Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., Korbel, J. O., Lander, E. S., Lee, C., Lehrach, H., Mardis, E. R., Marth, G. T., McVean, G. A., Nickerson, D. A., Schmidt, J. P., Sherry, S. T., Wang, J., Wilson, R. K., Gibbs, R. A., Dinh, H., Kovar, C., Lee, S., Lewis, L., Muzny, D., Reid, J., Wang, M., Wang, J., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Li, G., Li, J., Li, Y., Li, Z., Liu, X., Lu, Y., Ma, X., Su, Z., Tai, S., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Yin, Y., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Zhou, Y., Lander, E. S., Altshuler, D. M., Gabriel, S. B., Gupta, N., Flicek, P., Clarke, L., Leinonen, R., Smith, R. E., Zheng-Bradley, X., Bentley, D. R., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Lehrach, H., Sudbrak, R., Albrecht, M. W., Amstislavskiy, V. S., Borodina, T. A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.-L., Sherry, S. T., McVean, G. A., Mardis,

E. R., Wilson, R. K., Fulton, L., Fulton, R., Weinstock, G. M., Durbin, R. M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T. M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Schmidt, J. P., Davies, C. J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Gibbs, R. A., Yu, F., Bainbridge, M., Challis, D., Evani, U. S., Lu, J., Muzny, D., Nagaswamy, U., Reid, J., Sabo, A., Wang, Y., Yu, J., Wang, J., Coin, L. J. M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Y., Li, Z., Lin, H., Liu, B., Luo, R., Qin, N., Shao, H., Wang, B., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Marth, G. T., Garrison, E. P., Kural, D., Lee, W.-P., Fung Leong, W., Ward, A. N., Wu, J., Zhang, M., Lee, C., Griffin, L., Hsieh, C.-H., Mills, R. E., Shi, X., von Grotthuss, M., Zhang, C., Daly, M. J., DePristo, M. A., Altshuler, D. M., Banks, E., Bhatia, G., Carneiro, M. O., del Angel, G., Gabriel, S. B., Genovese, G., Gupta, N., Handsaker, R. E., Hartl, C., Lander, E. S., McCarroll, S. A., Nemesh, J. C., Poplin, R. E., Schaffner, S. F., Shakir, K., Yoon, S. C., Lihm, J., Makarov, V., Jin, H., Kim, W., Cheol Kim, K., Korbel, J. O., Rausch, T., Flicek, P., Beal, K., Clarke, L., Cunningham, F., Herrero, J., McLaren, W. M., Ritchie, G. R. S., Smith, R. E., Zheng-Bradley, X., Clark, A. G., Gottipati, S., Keinan, A., Rodriguez-Flores, J. L., Sabeti, P. C., Grossman, S. R., Tabrizi, S., Tariyal, R., Cooper, D. N., Ball, E. V., Stenson, P. D., Bentley, D. R., Barnes, B., Bauer, M., Keira Cheetham, R., Cox, T., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J., Shaw, R., Ye, K., Batzer, M. A., Konkel, M. K., Walker, J. A., MacArthur, D. G., Lek, M., Sudbrak, Amstislavskiy, V. S., Herwig, R., Shriver, M. D., Bustamante, C. D., Byrnes, J. K., De La Vega, F. M., Gravel, S., Kenny, E. E., Kidd, J. M., Lacroute, P., Maples, B. K., Moreno-Estrada, A., Zakharia, F., Halperin, E., Baran, Y., Craig, D. W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A. A., Sinari, S. A., Squire, K., Sherry, S. T., Xiao, C., Sebat, J., Bafna, V., Ye, K., Burchard, E. G., Hernandez, R. D., Gignoux, C. R., Haussler, D., Katzman, S. J., James Kent, W., Howie, B., & Ruiz-Linares, A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature, 491*(7422), 56-65. https://doi.org/10.1038/nature11632

Melby, J. N., & Conger, R. D. (2001). The Iowa Family Interaction Rating Scales: Instrument summary. In P. K. Kerig & K. M. Lindahl (Eds.), *Family observational coding systems: Resources for systemic research* (pp. 33–58). Erlbaum.

Micalizzi, L., Ronald, A., & Saudino, K. J. (2016). A genetically informed cross-lagged analysis of autistic-like traits and affective problems in early childhood. *Journal of Abnormal Child Psychology, 44*(5), 937-947. https://doi.org/10.1007/s10802-015-0088-6

Micalizzi, L., Wang, M., & Saudino, K. J. (2017). Difficult temperament and negative parenting in early childhood: A genetically informed cross-lagged analysis. *Developmental Science, 20*(2). https://doi.org/10.1111/desc.12355

Molenaar, D., & Dolan, C. V. (2014). Testing systematic genotype by environment interactions using item level data. *Behavior Genetics, 44*(3), 212-231. https://doi.org/10.1007/s10519-014-9647-9

Montgomery, D. E., & Koeltzow, T. E. (2010). A review of the day–night task: The Stroop paradigm and interference control in young children. *Developmental Review, 30*, 308-330. https://doi.org/10.1016/j.dr.2010.07.001

Mook-Kanamori, D. O., van Beijsterveldt, C. E., Steegers, E. A., Aulchenko, Y. S., Raat, H., Hofman, A., Eilers, P. H., Boomsma, D. I., & Jaddoe, V. W. (2012). Heritability estimates of body size in fetal life and early childhood. *PLoS One, 7*(7), e39901. https://doi.org/10.1371/journal.pone.0039901

Morris, T. T., Davies, N. M., Hemani, G., & Smith, G. D. (2019). Why are education, socioeconomic position and intelligence genetically correlated? *bioRxiv*, 630426. https://doi.org/10.1101/630426

NCD Risk Factor Collaboration. (2016). A century of trends in adult human height. *Elife, 5*. https://doi.org/10.7554/eLife.13410

Neale, M. C., & Stevenson, J. (1989). Rater bias in the EASI temperament scales: A twin study. *Journal of Personality and Social Psychology, 56*(3), 446-455. https://doi.org/10.1037//0022-3514.56.3.446

Neiss, M., & Rowe, D. C. (2000). Parental education and child's verbal IQ in adoptive and biological families in the National Longitudinal Study of Adolescent Health. *Behavior Genetetics, 30*, 487-495. https://doi.org/10.1023/a:1010254918997

Nelson, C. A., 3rd, Zeanah, C. H., & Fox, N. A. (2019). How early experience shapes human development: The case of psychosocial deprivation. *Neural Plasticity, 2019*. https://doi.org/https://doi.org/10.1155/2019/1676285

Nguyen, B. H., Perusse, D., Paquet, J., Petit, D., Boivin, M., Tremblay, R. E., & Montplaisir, J. (2008). Sleep terrors in children: A prospective study of twins. *Pediatrics, 122*(6), e1164-1167. https://doi.org/10.1542/peds.2008-1303

Nichols, P. L., & Broman, S. H. (1974). Familial resemblance in infant mental development. *Developmental Psychology, 10*(3), 442-446. https://doi.org/10.1037/h0036430

Niklas, F., & Schneider, W. (2013). Home literacy environment and the beginning of reading and spelling. *Contemporary Educational Psychology, 31*(1), 40-50.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-664. https://doi.org/10.1521/soco.19.6.625.20886

Nunnally, J. C. (1967). *Psychometric theory*. McGraw-Hill.

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlskog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., Porteous, D. J., Freese, J., Herd, P., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., Hinds, D. A., Huber, K. E., Kleinman, A., Litterman, N. K., McCreight, J. C., McIntyre, M. H., Mountain, J. L., Northover, C. A. M., Pitts, S. J., Sathirapongsasuti, J. F., Sazonova, O. V., Shelton, J. F., Shringarpure, S., Tung, J. Y., Vacic, V., Wilson, C. H., Fontana, M. A., Pers, T. H., Rietveld, C. A., Chen, G.-B., Emilsson, V., Meddens, S. F. W., Pickrell, J. K., Thom, K., Timshel, P., de Vlaming, R., Abdellaoui, A., Ahluwalia, T. S., Bacelis, J., Baumbach, C., Bjornsdottir, G., Brandsma, J. H., Concas, M. P., Derringer, J., Galesloot, T. E., Girotto, G., Gupta, R., Hall, L. M., Harris, S. E., Hofer, E., Horikoshi, M., Huffman, J. E., Kaasik, K., Kalafati, I. P., Karlsson, R., Lahti, J., van der Lee, S. J., de Leeuw, C., Lind, P. A., Lindgren, K.-O., Liu, T., Mangino, M., Marten, J., Mihailov, E., Miller, M. B., van der Most, P. J., Oldmeadow, C., Payton, A., Pervjakova, N., Peyrot, W. J., Qian, Y., Raitakari, O., Rueedi, R., Salvi, E., Schmidt, B., Schraut, K. E., Shi, J., Smith, A. V., Poot, R. A., Pourcain, B. S., Teumer, A., Thorleifsson, G., Verweij, N., Vuckovic, D., Wellmann, J., Westra, H.-J., Yang, J., Zhao, W., Zhu, Z., Alizadeh, B. Z., Amin, N., Bakshi, A., Baumeister, S. E., Biino, G., Bønnelykke, K., Boyle, P. A., Campbell, H., Cappuccio, F. P., Davies, G., De Neve, J.-E., Deloukas, P., Demuth, I., Ding, J., Eibich, P., Eisele, L., Eklund, N., Evans, D. M., Faul, J. D., Feitosa, M. F., Forstner, A. J., Gandin, I., Gunnarsson, B., Halldórsson, B. V., Harris, T. B., Heath, A. C., Hocking, L. J., Holliday, E. G., Homuth, G., Horan, M. A., Hottenga, J.-J., de Jager, P. L., Joshi, P. K., Jugessur, A., Kaakinen, M. A., Kähönen, M., Kanoni, S., Keltigangas-Järvinen, L., Kiemeney, L. A. L. M., Kolcic, I., Koskinen, S., Kraja, A. T., Kroh, M., Kutalik, Z., Latvala, A., Launer, L. J., Lebreton, M. P., Levinson, D. F., Lichtenstein, P., Lichtner, P., Liewald, D. C. M., Loukola, A., Madden, P. A., Mägi, R., Mäki-Opas, T., Marioni, R. E., Marques-Vidal, P., Meddens, G. A., McMahon, G., Meisinger, C., Meitinger, T., Milaneschi, Y., Milani, L., Montgomery, G. W., Myhre, R., Nelson, C. P., Nyholt, D. R., Ollier, W. E. R., Palotie, A., Paternoster, L., Pedersen, N. L., Petrovic, K. E., Räikkönen, K., Ring, S. M., Robino, A., Rostapshova, O., Rudan, I., Rustichini, A., Salomaa, V., Sanders, A. R., Sarin, A.-P., Schmidt, H., Scott, R. J., Smith, B. H., Smith, J. A., Staessen, J. A., Steinhagen-Thiessen, E., Strauch, K., Terracciano, A., Tobin, M. D., Ulivi, S., Vaccargiu, S., Quaye, L., van Rooij, F. J. A., Venturini, C., Vinkhuyzen, A. A. E., Völker, U., Völzke, H., Vonk, J. M., Vozzi, D., Waage, J., Ware, E. B., Willemsen, G., Attia, J. R., Bennett, D. A., Berger, K., Bertram, L., Bisgaard, H., Boomsma, D. I., Borecki, I. B., Bültmann, U., Chabris, C. F., Cucca, F., Cusi, D., Deary, I. J., Dedoussis, G. V., van Duijn, C. M., Eriksson, J. G., Franke, B., Franke, L., Gasparini, P., Gejman, P. V., Gieger, C., Grabe, H.-J., Gratten, J., Groenen, P. J. F., Gudnason, V., van der Harst, P., Hoffmann, W., Hyppönen, E., Iacono, W. G., Jacobsson, B., Järvelin, M.-R., Jöckel, K.-H., Kaprio, J., Kardia, S. L. R., Lehtimäki, T., Lehrer, S. F., Martin, N. G., McGue, M., Metspalu, A., Pendleton, N., Penninx, B. W. J. H., Perola, M., Pirastu, N., Pirastu, M., Polasek, O., Posthuma, D., Power, C., Province, M. A., Samani, N. J., Schlessinger, D., Schmidt, R., Sørensen, T. I. A., Spector, T. D., Stefansson, K., Thorsteinsdottir, U., Thurik, A. R., Timpson, N. J., Tiemeier, H., Uitterlinden, A. G., Vitart, V., Vollenweider, P., Weir, D. R., Wilson, J. F., Wright, A. F., Conley, D. C., Krueger, R. F., Smith, G. D., Hofman, A., Laibson, D. I., Medland, S. E., Yang, J., Esko, T., Watson, C., Jala, J., Conley, D., Koellinger, P. D., Johannesson, M., Laibson, D., Meyer, M. N., Lee, J. J., Kong, A., Yengo, L., Cesarini, D., Turley, P., Visscher, P. M., Beauchamp, J. P., andMe Research, T., Social Science Genetic Association, C., & LifeLines Cohort, S. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics, 54*(4), 437-449. https://doi.org/10.1038/s41588-022-01016-z

Orekhova, E. V., Stroganova, T. A., Posikera, I. N., & Malykh, S. B. (2003). Heritability and "environmentability" of electroencephalogram in infants:

The twin study. *Psychophysiology, 40*(5), 727-741. https://doi.org/10.1111/1469-8986.00073

Ouellet-Morin, I., Boivin, M., Dionne, G., Lupien, S. J., Arseneault, L., Barr, R. G., Perusse, D., & Tremblay, R. E. (2008). Variations in heritability of cortisol reactivity to stress as a function of early familial adversity among 19-month-old twins. *Archives of General Psychiatry, 65*(2), 211-218. https://doi.org/10.1001/archgenpsychiatry.2007.27

Ouellet-Morin, I., Dionne, G., Perusse, D., Lupien, S. J., Arseneault, L., Barr, R. G., Tremblay, R. E., & Boivin, M. (2009). Daytime cortisol secretion in 6-month-old twins: Genetic and environmental contributions as a function of early familial adversity. *Biological Psychiatry, 65*(5), 409-416. https://doi.org/10.1016/j.biopsych.2008.10.003

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal, 372*. https://doi.org/https://doi.org/10.1136/bmj.n71

Page, M. J., Sterne, J. A. C., Higgins, J. P. T., & Egger, M. (2021, Mar). Investigating and dealing with publication bias and other reporting biases in meta-analyses of health research: A review. *Research Synthesis Methods, 12*(2), 248-259. https://doi.org/10.1002/jrsm.1468

Paltiel, L., Anita, H., Skjerden, T., Harbak, K., Bækken, S., Nina Kristin, S., Knudsen, G. P., & Magnus, P. (2014). The biobank of the Norwegian Mother and Child Cohort Study – present status. *Norsk Epidemiologi, 24*(1-2). https://doi.org/10.5324/nje.v24i1-2.1755

Papageorgiou, K. A., & Ronald, A. (2017). The genetic basis of psychological traits in infancy: Implications for understanding the causes of developmental psychopathology. In L. C. Centifanti & D. M. Williams (Eds.), *The Wiley handbook of developmental psychopathology* (pp. 235–258). Wiley Blackwell. https://doi.org/https://doi.org/10.1002/9781118554470.ch11

Pappa, I., Szekely, E., Mileva-Seitz, V. R., Luijk, M. P., Bakermans-Kranenburg, M. J., van, I. M. H., & Tiemeier, H. (2015). Beyond the usual suspects: A multidimensional genetic exploration of infant attachment disorganization and security. *Attachment and Human Development, 17*(3), 288–301. https://doi.org/https://doi.org/10.1080/14616734.2015.1037316

Pardini, D. A., Fite, P. J., & Burke, J. D. (2008). Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: The moderating effect of age and african-american ethnicity. *Journal of Abnormal Child Psychology, 36*(5), 647-662. https://doi.org/10.1007/s10802-007-9162-z

Peter, I., Vainder, M., & Livshits, G. (1999). Genetic analysis of motor milestones attainment in early childhood. *Twin Research, 2*(1), 1-9. https://doi.org/10.1375/136905299320566040

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C.-Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A. C., Kalungi, A., Koen, N., Majara, L., Schwarz, E., Smoller, J. W., Stahl, E. A., Sullivan, P. F., Vassos, E., Mowry, B., Prieto, M. L., Cuellar-Barboza, A., Bigdeli, T. B., Edenberg, H. J., Huang, H., & Duncan, L. E. (2019). Genome-wide association studies in ancestrally diverse populations: Opportunities, methods, pitfalls, and recommendations. *Cell, 179*(3), 589-603. https://doi.org/10.1016/j.cell.2019.08.051

Petitclerc, A., Boivin, M., Dionne, G., Perusse, D., & Tremblay, R. E. (2011). Genetic and environmental etiology of disregard for rules. *Behavior Genetics, 41*(2), 192-200. https://doi.org/10.1007/s10519-010-9393-6

Petrill, S. A., Pike, A., Price, T., & Plomin, R. (2004). Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. *Intelligence, 32*, 445–460.

Phares, V., Lopez, E., Fields, S., Kamboukos, D., & Duhig, A. M. (2005). Are fathers involved in pediatric psychology research and treatment? *Journal of Pediatric Psychology, 30*(8), 631-643. https://doi.org/10.1093/jpepsy/jsi050

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science, 10*(3), 282-306. https://doi.org/10.1177/1745691615577701

Pike, A., McGuire, S., Hetherington, E., Reiss, D., & Plomin, R. (1996). Family environment and adolescent depressive symptoms and antisocial behavior: A multivariate genetic analysis. *Developmental Psychology, 32*(4), 590-603.

Pimpin, L., Ambrosini, G. L., Llewellyn, C. H., Johnson, L., van Jaarsveld, C. H., Jebb, S. A., & Wardle, J. (2013). Dietary intake of young twins: Nature or nurture? *The American Journal of Clinical Nutrition, 98*(5), 1326-1334. https://doi.org/10.3945/ajcn.113.065250

Pingault, J.-B., Allegrini, A. G., Odigie, T., Frach, L., Baldwin, J. R., Rijsdijk, F., & Dudbridge, F. (2022). Research review: How to interpret associations between polygenic scores, environmental risks, and phenotypes. *Journal of Child Psychology and Psychiatry, 63*(10), 1125-1139. https://doi.org/https://doi.org/10.1111/jcpp.13607

Planalp, E. M., Van Hulle, C., Lemery-Chalfant, K., & Goldsmith, H. H. (2017). Genetic and environmental contributions to the development of positive affect in infancy. *Emotion, 17*(3), 412-420. https://doi.org/10.1037/emo0000238

Plomin, R. (1986). Multivariate analysis and developmental behavioral genetics: Developmental change as well as continuity. *Behavior Genetics, 16*, 25-43. https://doi.org/10.1007/BF01065477

Plomin, R. (2011). Commentary: Why are children in the same family so different? Non-shared environment three decades later. *International Journal of Epidemiology, 40*(3), 582-592. https://doi.org/10.1093/ije/dyq144

Plomin, R. (2018). *Blueprint: How DNA makes us who we are*. Allen Lane.

Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on “environmental” measures. *Behavioral and Brain Sciences, 14*(3), 373-386. https://doi.org/10.1017/S0140525X00070278

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10*(1), 1-16. https://doi.org/10.1017/S0140525X00055941

Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry, 20*(1), 98-108. https://doi.org/10.1038/mp.2014.105

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science, 11*(1), 3-23. https://doi.org/10.1177/1745691615617439

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84*(2), 309-322.

Plomin, R., Emde, R. N., Braungart, J. M., Campos, J., Corley, R., Fulker, D. W., Kagan, J., Reznick, J. S., Robinson, J., Zahn-Waxler, C., & et al. (1993). Genetic change and continuity from fourteen to twenty months: The MacArthur Longitudinal Twin Study. *Child Development, 64*(5), 1354-1376. https://doi.org/10.2307/1131539

Plomin, R., Fulker, D. W., Corley, R. P., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 Years: A parent-offspring adoption study. *Psychological Science, 8*(6), 442-447. https://doi.org/10.1111/j.1467-9280.1997.tb00458.x

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*, 592-617. https://doi.org/10.1037/0033-2909.131.4.592

Plomin, R., & Rowe, D. C. (1979). Genetic and environmental etiology of social behavior in infancy. *Developmental Psychology, 15*(1), 62-72. https://doi.org/10.1037/h0078078

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

Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics, 47*(7), 702-709. https://doi.org/10.1038/ng.3285

Polderman, T. J. C., Posthuma, D., De Sonneville, L. M., Stins, J. F., Verhulst, F. C., & Boomsma, D. I. (2007). Genetic analyses of the stability of executive functioning during childhood. *Biological Psychology, 76*, 11-20. https://doi.org/10.1016/j.biopsycho.2007.05.002

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

Price, T. S., Eley, T. C., Dale, P. S., Stevenson, J., Saudino, K., & Plomin, R. (2000). Genetic and environmental covariation between verbal and nonverbal cognitive development in infancy. *Child Development, 71*(4), 948-959. https://doi.org/10.1111/1467-8624.00201

Price, T. S., Simonoff, E., Asherson, P., Curran, S., Kuntsi, J., Waldman, I., & Plomin, R. (2005). Continuity and change in preschool ADHD symptoms: Longitudinal genetic analysis with contrast effects. *Behavior Genetetics, 35*(2), 121-132. https://doi.org/10.1007/s10519-004-1013-x

Prior, M., Bavin, E., & Ong, B. (2011). Predictors of school readiness in five- to six-year-old children from an Australian longitudinal community sample. *Educational Psychology, 31*(1), 3-16. https://doi.org/10.1080/01443410.2010.541048

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics, 81*(3), 559-575. https://doi.org/10.1086/519795

Pushina, N. N., Orekhova, E. V., & Stroganova, T. A. (2005). Age-related and individual differences in the performance of a delayed response task (the A-not-B task) in infant twins aged 7-12 months. *Neuroscience and Behavioral Physiology, 35*(5), 481-490. https://doi.org/10.1007/s11055-005-0083-4

R Core Team. (2020). *A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rescorla, L., & Alley, A. (2001). Validation of the language development survey (LDS): A parent report tool for identifying language delay in toddlers. *Journal of Speech, Language, and Hearing Research, 44*, 434-445. https://doi.org/10.1044/1092-4388(2001/035)

Revelle, W. (2022). *psych: Procedures for personality and psychological research*. Northwestern University. https://CRAN.R-project.org/package=psych

Reznick, J. S., Corley, R., & Robinson, J. (1997). A longitudinal twin study of intelligence in the second year. *Monographs of the Society for Research in Child Development, 62*(1), i-vi, 1-154; discussion 155-160. https://doi.org/10.2307/1166193

Rhea, S.-A., Bricker, J. B., Wadsworth, S. J., & Corley, R. P. (2013). The Colorado Adoption Project. *Twin Research and Human Genetics, 16*(1), 358-365. https://doi.org/10.1017/thg.2012.109

Rhee, S. H., Corley, R. P., Friedman, N. P., Hewitt, J. K., Hink, L. K., Johnson, D. P., Robinson, J., Smith, A. K., & Young, S. E. (2012). The etiology of observed negative emotionality from 14 to 24 months. *Frontiers in Genetics, 3*, 9. https://doi.org/10.3389/fgene.2012.00009

Rhee, S. H., Cosgrove, V. E., Schmitz, S., Haberstick, B. C., Corley, R. C., & Hewitt, J. K. (2007). Early childhood temperament and the covariation between internalizing and externalizing behavior in school-aged children. *Twin Research and Human Genetics, 10*(1), 33-44. https://doi.org/10.1375/twin.10.1.33

Rhee, S. H., Friedman, N. P., Boeldt, D. L., Corley, R. P., Hewitt, J. K., Knafo, A., Lahey, B. B., Robinson, J., Van Hulle, C. A., Waldman, I. D., Young,

S. E., & Zahn-Waxler, C. (2013). Early concern and disregard for others as predictors of antisocial behavior. *Journal of Child Psychology and Psychiatry, 54*(2), 157-166. https://doi.org/10.1111/j.1469-7610.2012.02574.x

Rhee, S. H., Friedman, N. P., Corley, R. P., Hewitt, J. K., Hink, L. K., Johnson, D. P., Smith Watts, A. K., Young, S. E., Robinson, J., Waldman, I. D., & Zahn-Waxler, C. (2016). An examination of the developmental propensity model of conduct problems. *Journal of Abnormal Psychology, 125*(4), 550-564. https://doi.org/10.1037/abn0000128

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., Lewis, G., Harold, G. T., & Thapar, A. (2013). Examining the role of passive gene–environment correlation in childhood depression using a novel genetically sensitive design. *Development and Psychopathology, 25*(1), 37-50. https://doi.org/10.1017/S0954579412000880

Rice, M. L., Wilcox, K. A., Liebhaber, G. K., & Hadley, P. A. (1989). *The speech and Language Assessment Scale*. University of Kansas, USA.

Rice, M. L., Zubrick, S. R., Taylor, C. L., Gayan, J., & Bontempo, D. E. (2014). Late language emergence in 24-month-old twins: Heritable and increased risk for late language emergence in twins. *Journal of Speech, Language, and Hearing Research, 57*(3), 917-928. https://doi.org/10.1044/1092-4388(2013/12-0350)

Riese, M. L. (1990a). Genetic influences on neonatal temperament. *Acta Geneticae Medicae et Gemellologiae, 39*(2), 207-213. https://doi.org/10.1017/s0001566000005432

Riese, M. L. (1990b). Neonatal temperament in monozygotic and dizygotic twin pairs. *Child Development, 61*(4), 1230-1237. https://doi.org/10.2307/1130890

Robinson, J. L., Kagan, J., Reznick, J. S., & Corley, R. (1992). The heritability of inhibited and uninhibited behavior: A twin study. *Developmental Psychology, 28*(6), 1030-1037. https://doi.org/10.1037/0012-1649.28.6.1030

Robinson, M. R., Kleinman, A., Graff, M., Vinkhuyzen, A. A. E., Couper, D., Miller, M. B., Peyrot, W. J., Abdellaoui, A., Zietsch, B. P., Nolte, I. M., van Vliet-Ostaptchouk, J. V., Snieder, H., Alizadeh, B. Z., Boezen, H. M., Franke, L., van der Harst, P., Navis, G., Rots, M., Snieder, H., Swertz, M., Wolffenbuttel, B. H. R., Wijmenga, C., Abecasis, G. R., Absher, D., Alavere, H., Albrecht, E., Allen, H. L., Almgren, P., Amin, N., Amouyel, P., Anderson, D., Arnold, A. M., Arveiler, D., Aspelund, T., Asselbergs, F. W., Assimes, T. L., Atalay, M., Attwood, A. P., Atwood, L. D., Bakker, S. J. L., Balkau, B., Balmforth, A. J., Barlassina, C., Barroso, I., Basart, H., Bauer, S., Beckmann, J. S., Beilby, J. P., Bennett, A. J., Ben-Shlomo, Y., Bergman, R. N., Bergmann, S., Berndt, S. I., Biffar, R., Di Blasio, A. M., Boehm, B. O., Boehnke, M., Boeing, H., Boerwinkle, E., Bolton, J. L., Bonnefond, A., Bonnycastle, L. L., Boomsma, D. I., Borecki, I. B., Bornstein, S. R., Bouatia-Naji, N., Boucher, G., Bragg-Gresham, J. L., Brambilla, P., Bruinenberg, M., Buchanan, T. A., Buechler, C., Cadby, G., Campbell, H., Caulfield, M. J., Cavalcanti-Proença, C., Cesana, G., Chanock, S. J., Chasman, D. I., Chen, Y. I., Chines, P. S., Clegg, D. J., Coin, L., Collins, F. S., Connell, J. M., Cookson, W., Cooper, M. N., Croteau-Chonka, D. C., Cupples, L. A., Cusi, D., Day, F. R., Day, I. N. M., Dedoussis, G. V., Dei, M., Deloukas, P., Dermitzakis, E. T., Dimas, A. S., Dimitriou, M., Dixon, A. L., Dörr, M., van Duijn, C. M., Ebrahim, S., Edkins, S., Eiriksdottir, G., Eisinger, K., Eklund, N., Elliott, P., Erbel, R., Erdmann, J., Erdos, M. R., Eriksson, J. G., Esko, T., Estrada, K., Evans, D. M., de Faire, U., Fall, T., Farrall, M., Feitosa, M. F., Ferrario, M. M., Ferreira, T., Ferrières, J., Fischer, K., Fisher, E., Fowkes, G., Fox, C. S., Franke, L., Franks, P. W., Fraser, R. M., Frau, F., Frayling, T., Freimer, N. B., Froguel, P., Fu, M., Gaget, S., Ganna, A., Gejman, P. V., Gentilini, D., Geus, E. J. C., Gieger, C., Gigante, B., Gjesing, A. P., Glazer, N. L., Goddard, M. E., Goel, A., Grallert, H., Gräßler, J., Grönberg, H., Groop, L. C., Groves, C. J., Gudnason, V., Guiducci, C., Gustafsson, S., Gyllensten, U., Hall, A. S., Hall, P., Hallmans, G., Hamsten, A., Hansen, T., Haritunians, T., Harris, T. B., van der Harst, P., Hartikainen, A., Hassanali, N., Hattersley, A. T., Havulinna, A. S., Hayward, C., Heard-Costa, N. L., Heath, A. C., Hebebrand, J., Heid, I. M., den Heijer, M., Hengstenberg, C., Herzig, K., Hicks, A. A., Hingorani, A., Hinney, A., Hirschhorn, J. N., Hofman, A., Holmes, C. C., Homuth, G., Hottenga, J., Hovingh, K. G., Hu, F. B., Hu, Y., Huffman, J. E., Hui, J., Huikuri, H., Humphries, S. E., Hung, J., Hunt, S. E., Hunter, D., Hveem, K., Hyppönen, E., Igl, W., Illig, T., Ingelsson, E., Iribarren, C., Isomaa, B., Jackson, A. U., Jacobs, K. B., James, A. L., Jansson, J., Jarick, I., Jarvelin, M., Jöckel, K. H., Johansson, Å., Johnson, T., Jolley, J., Jørgensen, T., Jousilahti, P., Jula, A., Justice, A. E., Kaakinen, M., Kähönen, M., Kajantie, E., Kanoni, S., Kao, W. H. L., Kaplan, L. M., Kaplan, R. C., Kaprio, J., Kapur, K., Karpe, F., Kathiresan, S., Kee, F., Keinanen-Kiukaanniemi, S. M., Ketkar, S., Kettunen, J., Khaw, K., Kiemeney, L. A., Kilpeläinen, T. O., Kinnunen, L., Kivimaki, M., Kivmaki, M., Van der Klauw, M. M., Kleber, M. E., Knowles, J. W., Koenig, W., Kolcic, I., Kolovou, G., König, I. R., Koskinen, S., Kovacs, P., Kraft, P., Kraja, A. T., Kristiansson, K., Krjutškov, K., Kroemer, H. K., Krohn, J. P., Krzelj, V., Kuh, D., Kulzer, J. R., Kumari, M., Kutalik, Z., Kuulasmaa, K., Kuusisto, J., Kvaloy, K., Laakso, M., Laitinen, J. H., Lakka, T. A., Lamina, C., Langenberg, C., Lantieri, O., Lathrop, G. M., Launer, L. J., Lawlor, D. A., Lawrence, R. W., Leach, I. M., Lecoeur, C., Lee, S. H., Lehtimäki, T., Leitzmann, M. F., Lettre, G., Levinson, D. F., Li, G., Li, S., Liang, L., Lin, D., Lind, L., Lindgren, C. M., Lindström, J., Liu, J., Liuzzi, A., Locke, A. E., Lokki, M., Loley, C., Loos, R. J. F., Lorentzon, M., Luan, J., Luben, R. N., Ludwig, B., Madden, P. A., Mägi, R., Magnusson, P. K. E., Mangino, M., Manunta, P., Marek, D., Marre, M., Martin, N. G., The LifeLines Cohort Study, & Genetic Investigation of Anthropometric Traits

Consortium. (2017). Genetic evidence of assortative mating in humans. *Nature Human Behaviour, 1*(1), 0016. https://doi.org/10.1038/s41562-016-0016

Roisman, G. I., & Fraley, R. C. (2006). The limits of genetic influence: A behavior-genetic analysis of infant-caregiver relationship quality and temperament. *Child Development, 77*(6), 1656-1667. https://doi.org/10.1111/j.1467-8624.2006.00965.x

Roisman, G. I., & Fraley, R. C. (2008). A behavior-genetic study of parenting quality, infant attachment security, and their covariation in a nationally representative sample. *Developmental Psychology, 44*(3), 831-839. https://doi.org/10.1037/0012-1649.44.3.831

Ronald, A., Edelson, L. R., Asherson, P., & Saudino, K. J. (2010). Exploring the relationship between autistic-like traits and ADHD behaviors in early childhood: Findings from a community twin study of 2-year-olds. *Journal of Abnormal Child Psychology, 38*(2), 185-196. https://doi.org/10.1007/s10802-009-9366-5

Ronald, A., Happe, F., & Plomin, R. (2008). A twin study investigating the genetic and environmental aetiologies of parent, teacher and child ratings of autistic-like traits and their overlap. *European Child and Adolescent Psychiatry, 17*(8), 473–483. https://doi.org/https://doi.org/10.1007/s00787-008-0689-5

Ronalds, G. A., De Stavola, B. L., & Leon, D. A. (2005). The cognitive cost of being a twin: Evidence from comparisons within families in the Aberdeen children of the 1950s cohort study. *British Medical Journal, 331*(7528), 1306. https://doi.org/https://doi.org/10.1136/bmj.38633.594387.3A

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://doi.org/https://doi.org/10.18637/jss.v048.io2

Rutter, M. (1998, May). Developmental catch-up, and deficit, following adoption after severe global early privation. *The Journal of Child Psychology and Psychiatry, 39*(4), 465–476. https://www.ncbi.nlm.nih.gov/pubmed/9599775

Rutter, M. (2000). Psychosocial influences: Critiques, findings, and research needs. *Development and Psychopathology, 12*(3), 375-405. https://doi.org/10.1017/S0954579400003072

Saccuzzo, D. P., Johnson, N. E., & Russell, G. (1992). Verbal versus performance IQs for gifted African-American, Caucasian, Filipino, and Hispanic children. *Psychological Assessment, 4*(2), 239. https://doi.org/10.1037/1040-3590.4.2.239

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C. M., & Reichenberg, A. (2014). The familial risk of autism. *JAMA, 311*(17), 1770-1777. https://doi.org/10.1001/jama.2014.4144

Saudino, K. J. (2012). Sources of continuity and change in activity level in early childhood. *Child Development, 83*(1), 266-281. https://doi.org/10.1111/j.1467-8624.2011.01680.x

Saudino, K. J., Carter, A. S., Purper-Ouakil, D., & Gorwood, P. (2008). The etiology of behavioral problems and competencies in very young twins. *Journal of Abnormal Psychology, 117*(1), 48-62. https://doi.org/10.1037/0021-843x.117.1.48

Saudino, K. J., & Eaton, W. O. (1991). Infant temperament and genetics: An objective twin study of motor activity level. *Child Development, 62*(5), 1167-1174. https://doi.org/10.2307/1131160

Saudino, K. J., Plomin, R., & DeFries, J. C. (1996). Tester-rated temperament at 14, 20 and 24 months: Environmental change and genetic continuity.

British Journal of Developmental Psychology, 14(2), 129-144. https://doi.org/10.1111/j.2044-835X.1996.tb00697.x

Saudino, K. J., Ronald, A., & Plomin, R. (2005). The etiology of behavior problems in 7-year-old twins: Substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. *Journal of Abnormal Child Psychology, 33*(1), 113–130. https://doi.org/https://doi.org/10.1007/s10802-005-0939-7

Saudino, K. J., Wang, M., Flom, M., & Asherson, P. (2018). Genetic and environmental links between motor activity level and attention problems in early childhood. *Developmental Science, 21*(5), e12630. https://doi.org/10.1111/desc.12630

Saudino, K. J., & Zapfe, J. A. (2008). Genetic influences on activity level in early childhood: Do situations matter? *Child Development, 79*(4), 930-943. https://doi.org/10.1111/j.1467-8624.2008.01168.x

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*(2), 424-435. https://doi.org/10.2307/1129703

Scarr-Salapatek, S. (1971). Race, social class, and IQ. *Science, 174*(4016), 1285-1295. https://doi.org/10.1126/science.174.4016.1285

Schmitz, S., Fulker, D. W., Plomin, R., Zahn-Waxler, C., Emde, R. N., & DeFries, J. C. (1999). Temperament and problem behavior during early childhood. *International Journal of Behavioral Development, 23*(2), 333-355. https://doi.org/10.1080/016502599383856

Schoeler, T., Speed, D., Porcu, E., Pirastu, N., Pingault, J.-B., & Kutalik, Z. (2022). Correction for participation bias in the UK Biobank reveals non-negligible impact on genetic associations and downstream analyses. *bioRxiv*. https://doi.org/10.1101/2022.09.28.509845

Schumann, L., Boivin, M., Paquin, S., Lacourse, E., Brendgen, M., Vitaro, F., Dionne, G., Tremblay, R. E., & Booij, L. (2017). Persistence and innovation effects in genetic and environmental factors in negative emotionality during infancy: A twin study. *PLoS One, 12*(4), e0176601. https://doi.org/10.1371/journal.pone.0176601

Selzam, S., Dale, P. S., Wagner, R. K., DeFries, J. C., Cederlof, M., O'Reilly, P. F., Krapohl, E., & Plomin, R. (2017). Genome-wide polygenic scores predict reading performance throughout the school years. *Scientific Studies of Reading, 21*, 334-349. https://doi.org/10.1080/10888438.2017.1299152

Shelton, K. K., Frick, P. J., & Wootton, J. (1996a). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology, 25*(3), 317-329.

Shelton, K. K., Frick, P. J., & Wootton, J. M. (1996b). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology, 25*(3), 317-329. https://doi.org/10.1207/s15374424jccp2503_8

Silberg, J. L., Gillespie, N., Moore, A. A., Eaves, L. J., Bates, J., Aggen, S., Pfister, E., & Canino, G. (2015). Shared genetic and environmental influences on early temperament and preschool psychiatric disorders in Hispanic twins. *Twin Research and Human Genetics, 18*(2), 171-178. https://doi.org/10.1017/thg.2014.88

Silberg, J. L., Miguel, V. F., Murrelle, E. L., Prom, E., Bates, J. E., Canino, G., Egger, H., & Eaves, L. J. (2005). Genetic and environmental influences on temperament in the first year of life: The Puerto Rico Infant Twin Study (PRINTS). *Twin Research and Human Genetics, 8*(4), 328-336. https://doi.org/10.1375/1832427054936808

Silventoinen, K., Jelenkovic, A., Sund, R., Hur, Y. M., Yokoyama, Y., Honda, C., Hjelmborg, J., Moller, S., Ooki, S., Aaltonen, S., Ji, F., Ning, F., Pang, Z., Rebato, E., Busjahn, A., Kandler, C., Saudino, K. J., Jang, K. L., Cozen, W., Hwang, A. E., Mack, T. M., Gao, W., Yu, C., Li, L., Corley, R. P., Huibregtse, B. M., Christensen, K., Skytthe, A., Kyvik, K. O., Derom, C. A., Vlietinck, R. F., Loos, R. J., Heikkila, K., Wardle, J., Llewellyn, C. H., Fisher, A., McAdams, T. A., Eley, T. C., Gregory, A. M., He, M., Ding, X., Bjerregaard-Andersen, M., Beck-Nielsen, H., Sodemann, M., Tarnoki, A. D., Tarnoki, D. L., Stazi, M. A., Fagnani, C., D'Ippolito, C., Knafo-Noam, A., Mankuta, D., Abramson, L., Burt, S. A., Klump, K. L., Silberg, J. L., Eaves, L. J., Maes, H. H., Krueger, R. F., McGue, M., Pahlen, S., Gatz, M., Butler, D. A., Bartels, M., van Beijsterveldt, T. C., Craig, J. M., Saffery, R., Freitas, D. L., Maia, J. A., Dubois, L., Boivin, M., Brendgen, M., Dionne, G., Vitaro, F., Martin, N. G., Medland, S. E., Montgomery, G. W., Chong, Y., Swan, G. E., Krasnow, R., Magnusson, P. K., Pedersen, N. L., Tynelius, P., Lichtenstein, P., Haworth, C. M., Plomin, R., Bayasgalan, G., Narandalai, D., Harden, K. P., Tucker-Drob, E. M., Oncel, S. Y., Aliev, F., Spector, T., Mangino, M., Lachance, G., Baker, L. A., Tuvblad, C., Duncan, G. E., Buchwald, D., Willemsen, G., Rasmussen, F., Goldberg, J. H., Sorensen, T., Boomsma, D. I., & Kaprio, J. (2016). Genetic and environmental effects on body mass index from infancy to the onset of adulthood: an individual-based pooled analysis of 45 twin cohorts participating in the COllaborative project of Development of Anthropometrical measures in Twins (CODATwins) study. *The American Journal of Clinical Nutrition, 104*(2), 371-379. https://doi.org/10.3945/ajcn.116.130252

Silventoinen, K., Jelenkovic, A., Sund, R., Latvala, A., Honda, C., Inui, F., Tomizawa, R., Watanabe, M., Sakai, N., Rebato, E., Busjahn, A., Tyler, J., Hopper, J. L., Ordoñana, J. R., Sánchez-Romera, J. F., Colodro-Conde, L., Calais-Ferreira, L., Oliveira, V. C., Ferreira, P. H., Medda, E., Nisticò, L., Toccaceli, V., Derom, C. A., Vlietinck, R. F., Loos, R. J. F., Siribaddana, S. H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Duncan, G. E., Buchwald, D., Tynelius, P., Rasmussen, F., Tan, Q., Zhang, D., Pang, Z., Magnusson, P. K. E., Pedersen, N. L., Dahl Aslan, A. K., Hwang, A. E., Mack, T. M., Krueger, R. F., McGue, M., Pahlen, S., Brandt, I., Nilsen, T. S., Harris, J. R., Martin, N. G., Medland, S. E., Montgomery, G. W., Willemsen, G., Bartels, M., van Beijsterveldt, C. E. M., Franz, C. E., Kremen, W. S., Lyons, M. J., Silberg, J. L., Maes, H. H., Kandler, C., Nelson, T. L., Whitfield, K. E., Corley, R. P., Huibregtse, B. M., Gatz, M., Butler, D. A., Tarnoki, A. D., Tarnoki, D. L., Park, H. A., Lee, J., Lee, S. J., Sung, J., Yokoyama, Y., Sørensen, T. I. A., Boomsma, D. I., & Kaprio, J. (2020). Genetic and environmental variation in educational attainment: an individual-based analysis of 28 twin cohorts. *Scientific Reports, 10*(1), 12681. https://doi.org/10.1038/s41598-020-69526-6

Silventoinen, K., Kaprio, J., Dunkel, L., & Yokoyama, Y. (2012). Genetic and environmental influences on chest circumference during infancy: A longitudinal study of japanese twins. *Paediatric and Perinatal Epidemiology, 26*(6), 553-560. https://doi.org/10.1111/ppe.12003

Silventoinen, K., Kaprio, J., & Yokoyama, Y. (2011). Genetics of pre-pubertal growth: A longitudinal study of Japanese twins. *Annals of Human Biology, 38*(5), 608-614. https://doi.org/10.3109/03014460.2011.587453

Silventoinen, K., Karvonen, M., Sugimoto, M., Kaprio, J., Dunkel, L., & Yokoyama, Y. (2011). Genetics of head circumference in infancy: A longitudinal study of Japanese twins. *American Journal of Human Biology, 23*(5), 630-634. https://doi.org/10.1002/ajhb.21190

Silventoinen, K., Pietilainen, K. H., Tynelius, P., Sorensen, T. I., Kaprio, J., & Rasmussen, F. (2007). Genetic and environmental factors in relative weight from birth to age 18: The Swedish young male twins study. *International Journal of Obesity, 31*(4), 615-621. https://doi.org/10.1038/sj.ijo.0803577

Silventoinen, K., Pietilainen, K. H., Tynelius, P., Sorensen, T. I., Kaprio, J., & Rasmussen, F. (2008). Genetic regulation of growth from birth to 18 years of age: The Swedish young male twins study. *The American Journal of Human Biology, 20*(3), 292-298. https://doi.org/10.1002/ajhb.20717

Sim, F., Thompson, L., Marryat, L., Ramparsad, N., & Wilson, P. (2019). Predictive validity of preschool screening tools for language and behavioural difficulties: A PRISMA systematic review. *PLoS One, 14*, e0211409. https://doi.org/10.1371/journal.pone.0211409

Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell, 177*(1), 26-31. https://doi.org/10.1016/j.cell.2019.02.048

Smit, D. J., Luciano, M., Bartels, M., van Beijsterveldt, C. E., Wright, M. J., Hansell, N. K., Brunner, H. G., Estourgie-van Burk, G. F., de Geus, E. J., Martin, N. G., & Boomsma, D. I. (2010). Heritability of head size in Dutch and Australian twin families at ages 0-50 years. *Twin Research and Human Genetics, 13*(4), 370-380. https://doi.org/10.1375/twin.13.4.370

Smith, A. D., Herle, M., Fildes, A., Cooke, L., Steinsbekk, S., & Llewellyn, C. H. (2017). Food fussiness and food neophobia share a common etiology in early childhood. *Journal of Child Psychology and Psychiatry, 58*(2), 189-196. https://doi.org/10.1111/jcpp.12647

Smith, A. K., Rhee, S. H., Corley, R. P., Friedman, N. P., Hewitt, J. K., & Robinson, J. L. (2012). The magnitude of genetic and environmental influences on parental and observational measures of behavioral inhibition and shyness in toddlerhood. *Behavior Genetetics, 42*(5), 764-777. https://doi.org/10.1007/s10519-012-9551-0

Smith, L., van Jaarsveld, C. H. M., Llewellyn, C. H., Fildes, A., Lopez Sanchez, G. F., Wardle, J., & Fisher, A. (2017). Genetic and environmental influences on developmental milestones and movement: Results from the Gemini cohort study. *Research Quarterly for Exercise and Sport, 88*(4), 401-407. https://doi.org/10.1080/02701367.2017.1373268

Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2019). Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry, 24*(3), 409-420. https://doi.org/10.1038/s41380-017-0010-4

Soussignan, R., Boivin, M., Girard, A., Perusse, D., Liu, X., & Tremblay, R. E. (2009). Genetic and environmental etiology of emotional and social behaviors in 5-month-old infant twins: Influence of the social context. *Infant Behavior & Development, 32*(1), 1-9. https://doi.org/10.1016/j.infbeh.2008.09.002

Spinath, F. M., Ronald, A., Harlaar, N., Price, T. S., & Plomin, R. (2003). Phenotypic g early in life: On the etiology of general cognitive ability in a large population sample of twin children aged 2-4 years. *Intelligence, 31*(2), 195-210. https://doi.org/10.1016/S0160-2896(02)00110-1

Statistisk Sentralbyrå. (2001). Educational attainment of the population, 1 October 1999. https://www.ssb.no/en/utdanning/statistikker/utniv/arkiv/2001-03-07

Statistisk Sentralbyrå. (2009). *Educational attainment of the population, 1 October 2008*. https://www.ssb.no/en/utdanning/statistikker/utniv/arkiv/2009-08-25

Stevenson, J., & Fielding, J. (1985). Ratings of temperament in families of young twins. *British Journal of Developmental Psychology, 3*, 143–152. https://doi.org/10.1111/j.2044-835X.1985.tb00966.x

Stoolmiller, M. (1999). Implications of the restricted range of family environments for estimates of heritability and nonshared environment in behavior-genetic adoption studies. *Psychological Bulletin, 125*(4), 392-409. https://doi.org/10.1037/0033-2909.125.4.392

Stroganova, T., Tsetlin, M., Malykh, S., & Malakhovskaya, E. (2000). Biological principles of individual differences of children of the second half-year of life: Communication II. The nature of individual differences in temperamental features. *Human Physiology, 26*(3), 281-289.

Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language, 77*, 647-723. https://doi.org/10.1353/lan.2001.0247

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. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA, 283*(15), 2008-2012. https://doi.org/10.1001/jama.283.15.2008

Suizzo, M. A., & Stapleton, L. M. (2007, 2007/08/01). Home-based parental involvement in young children’s education: Examining the effects of maternal education across U.S. ethnic groups. *Educational Psychology, 27*(4), 533-556. https://doi.org/10.1080/01443410601159936

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics, 20*(8), 467-484. https://doi.org/10.1038/s41576-019-0127-1

Taylor, L. C., Clayton, J. D., & Rowley, S. J. (2004). Academic socialization: Understanding parental influences on children's school-related development in the early years. *Review of General Psychology, 8*(3), 163-178. https://doi.org/10.1037/1089-2680.8.3.163

The Norwegian Directorate of Education. (2020). *The Norwegian Education Mirror, 2019*. Retrieved 9th January from https://www.udir.no/in-english/education-mirror-2019/

The World Bank. (2023). *World Development Indicators*. Retrieved 9th January from https://datatopics.worldbank.org/world-development-indicators/

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

Tierney, N., Cook, D., McBain, M., & Fay, C. (2021). *naniar: Data structures, summaries, and visualisations for missing data. R package version 0.6.1.* https://CRAN.R-project.org/package=naniar

Touchette, E., Dionne, G., Forget-Dubois, N., Petit, D., Perusse, D., Falissard, B., Tremblay, R. E., Boivin, M., & Montplaisir, J. Y. (2013). Genetic and environmental influences on daytime and nighttime sleep duration in early childhood. *Pediatrics, 131*(6), e1874-1880. https://doi.org/10.1542/peds.2012-2284

Touwslager, R. N., Gielen, M., Derom, C., Mulder, A. L., Gerver, W. J., Zimmermann, L. J., Houben, A. J., Stehouwer, C. D., Vlietinck, R., Loos, R. J., & Zeegers, M. P. (2011). Determinants of infant growth in four age windows: A twin study. *The Journal of Pediatrics, 158*(4), 566-572.e562. https://doi.org/10.1016/j.jpeds.2010.10.005

Touwslager, R. N., Gielen, M., Mulder, A. L., Gerver, W. J., Zimmermann, L. J., Fowler, T., Houben, A. J., Stehouwer, C. D., Derom, C., Vlietinck, R., Loos, R. J., & Zeegers, M. P. (2011). Changes in genetic and environmental effects on growth during infancy. *The American Journal of Clinical Nutrition, 94*(6), 1568-1574. https://doi.org/10.3945/ajcn.111.012757

Tucker-Drob, E. M., & Bates, T. C. (2015). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138-149. https://doi.org/10.1177/0956797615612727

Tucker-Drob, E. M., & Briley, D. A. (2014). Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychological Bulletin, 140*, 949-979. https://doi.org/10.1037/a0035893

Tucker-Drob, E. M., & Harden, K. P. (2012). Early childhood cognitive development and parental cognitive stimulation: Evidence for reciprocal gene–environment transactions. *Developmental Science, 15*(2), 250-259. https://doi.org/10.1111/j.1467-7687.2011.01121.x

Tucker-Drob, E. M., Rhemtulla, M., Harden, K. P., Turkheimer, E., & Fask, D. (2011). Emergence of a Gene x socioeconomic status interaction on infant mental ability between 10 months and 2 years. *Psychological Science, 22*(1), 125-133. https://doi.org/10.1177/0956797610392926

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 9*(5), 160-164. https://doi.org/10.1111/1467-8721.00084

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, II. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science, 14*(6), 623-628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

Van den Noortgate, W., Lopez-Lopez, J. A., Marin-Martinez, F., & Sanchez-Meca, J. (2013, Jun). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45*(2), 576–594. https://doi.org/https://doi.org/10.3758/s13428-012-0261-6 van Dommelen, P., de Gunst, M. C., van der Vaart, A. W., & Boomsma, D. I. (2004). Genetic study of the height and weight process during infancy. *Twin Research, 7*(6), 607-616. https://doi.org/10.1375/1369052042663805

Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P. S., & St Pourcain, B. (2021). The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry, 62*(6), 728-738. https://doi.org/10.1111/jcpp.13327

Verhulst, B. (2017). A power calculator for the classical twin design. *Behavior Genetics, 47*(2), 255-261. https://doi.org/10.1007/s10519-016-9828-9

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

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

Wang, B., Baldwin, J. R., Schoeler, T., Cheesman, R., Barkhuizen, W., Dudbridge, F., Bann, D., Morris, T. T., & Pingault, J. B. (2021). Robust genetic nurture effects on education: A systematic review and meta-analysis based on 38,654 families across 8 cohorts. *American Journal of Human Genetics, 108*(9), 1780-1791. https://doi.org/10.1016/j.ajhg.2021.07.010

Wang, M., & Saudino, K. J. (2012). Genetic and environmental contributions to stability and change of sleep problems in toddlerhood. *Journal of Pediatric Psychology, 37*(6), 697-706. https://doi.org/10.1093/jpepsy/jss048

Wang, Y., Tsuo, K., Kanai, M., Neale, B. M., & Martin, A. R. (2022). Challenges and opportunities for developing more generalizable polygenic risk scores. *Annual Review of Biomedical Data Science, Advance online publication*. https://doi.org/10.1146/annurev-biodatasci-111721-074830

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. Psychological Corporation.

Wechsler, D. (2002). *WPPSI-III administration and scoring manual*. Psychological Corporation.

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

Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Houts, R., Matthews, T., Prinz, J. A., Richmond-Rakerd, L. S., Sugden, K., Williams, B., & Caspi, A. (2020). Using DNA from mothers and children to study parental investment in children's educational attainment. *Child Development, 91*(5), 1745-1761. https://doi.org/10.1111/cdev.13329

Whitfield, J. B., Treloar, S. A., Zhu, G., & Martin, N. G. (2001). Genetic and non-genetic factors affecting birthweight and adult body mass index. *Twin Research, 4*, 365-370. https://doi.org/10.1375/1369052012533

Wilson, R. S. (1972). Twins: Early mental development. *Science, 175*(4024), 914-917. https://doi.org/10.1126/science.175.4024.914

Wilson, R. S. (1974). Twins: Mental development in the preschool years. *Developmental Psychology, 10*(4), 580-588. https://doi.org/doi.org/10.1037/h0036596

Wilson, R. S. (1978). Synchronies in mental development: An epigenetic perspective. *Science, 202*(4371), 939-948. https://doi.org/10.1126/science.568822

Wilson, R. S. (1983). The Louisville Twin Study: Developmental synchronies in behavior. *Child Development, 54*(2), 298-316. https://doi.org/10.2307/1129693

Wilson, R. S. (1984). Twins and chronogenetics: Correlated pathways of development. *Acta Geneticae Medicae et Gemellologiae, 33*(2), 149-157. https://doi.org/10.1017/s0001566000007170

Wilson, R. S., & Harpring, E. B. (1972). Mental and motor development in infant twins. *Developmental Psychology, 7*(3), 277-287. https://doi.org/10.1037/h0033357

Wilson, R. S., & Matheny, A. P., Jr. (1976). Retardation and twin concordance in infant mental development: A reassessment. *Behavior Genetetics, 6*(3), 353-358. https://doi.org/10.1007/BF01065730

Wolchik, S. A., Tein, J. Y., Sandler, I. N., & Kim, H. J. (2016). Developmental cascade models of a parenting-focused program for divorced families on mental health problems and substance use in emerging adulthood. *Development and Psychopathology, 28*, 869-888. https://doi.org/10.1017/S0954579416000365

Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson tests of achievement (WJ III®)*. Riverside Publishing.

Woodward, K. E., Boeldt, D. L., Corley, R. P., DiLalla, L., Friedman, N. P., Hewitt, J. K., Mullineaux, P. Y., Robinson, J., & Rhee, S. H. (2018). Correlates of Positive Parenting Behaviors. *Behavior Genetics*. https://doi.org/10.1007/s10519-018-9906-2

World Health Organization. (2007). *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY*. World Health Organization.

Wright, S. (1921). Systems of mating. *Genetics, 6*(2), 111-178. https://doi.org/10.1093/genetics/6.2.167

Yew, S. G., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry, 54*, 516-524. https://doi.org/10.1111/jcpp.12009

Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. E. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence, 69*, 8-15. https://doi.org/10.1016/j.intell.2018.03.011

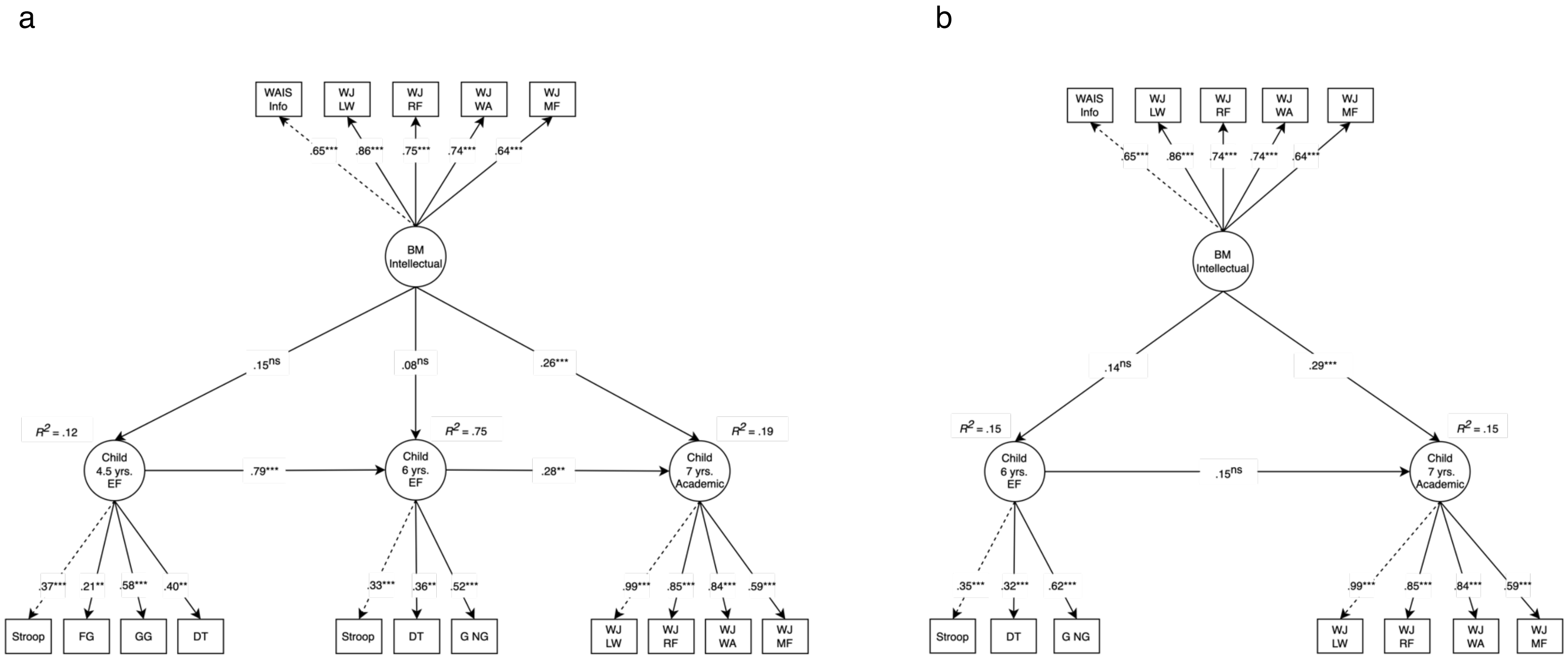
Zilanawala, A., Kelly, Y., & Sacker, A. (2016). Ethnic differences in longitudinal latent verbal profiles in the millennium cohort study. *European Journal of Public Health, 26*(6), 1011-1016. https://doi.org/10.1093/eurpub/ckw184

Appendix B

B.1 Sensitivity Analysis: Birth Parent Age

As not all the birth parent measures of intellectual performance were age-normed, I conducted a sensitivity analysis, examining whether birth parent age confounded associations between birth parents and children. I began by examining whether birth parents' age (when they were administered the measures of intellectual performance) was associated with their intellectual performance. There was no association between birth mother age and intellectual performance (β = -.004, 95% CI [-.12, 0.11], p = .95). However, there was a negative association between birth father age and intellectual performance (β = –.29, 95% CI [–.47, –0.11], p = .007). Consequently, I recomputed the main birth father and EF and language models, to test whether birth father age confounded any of the associations between birth parent intellectual performance and child EF, language, and academic test performance. Model fit declined from good, in the original models (RMSEA = .03–05, SRMR = .07–.08) to poor, when birth father age was added to the models (RMSEA = .08–.16, SRMR = .06–.16). There were no meaningful changes in associations between birth fathers and children, apart from that the effect estimate for the association between birth father intellectual performance and child language at 4.5 years reduced from β = .37 to β = .23.

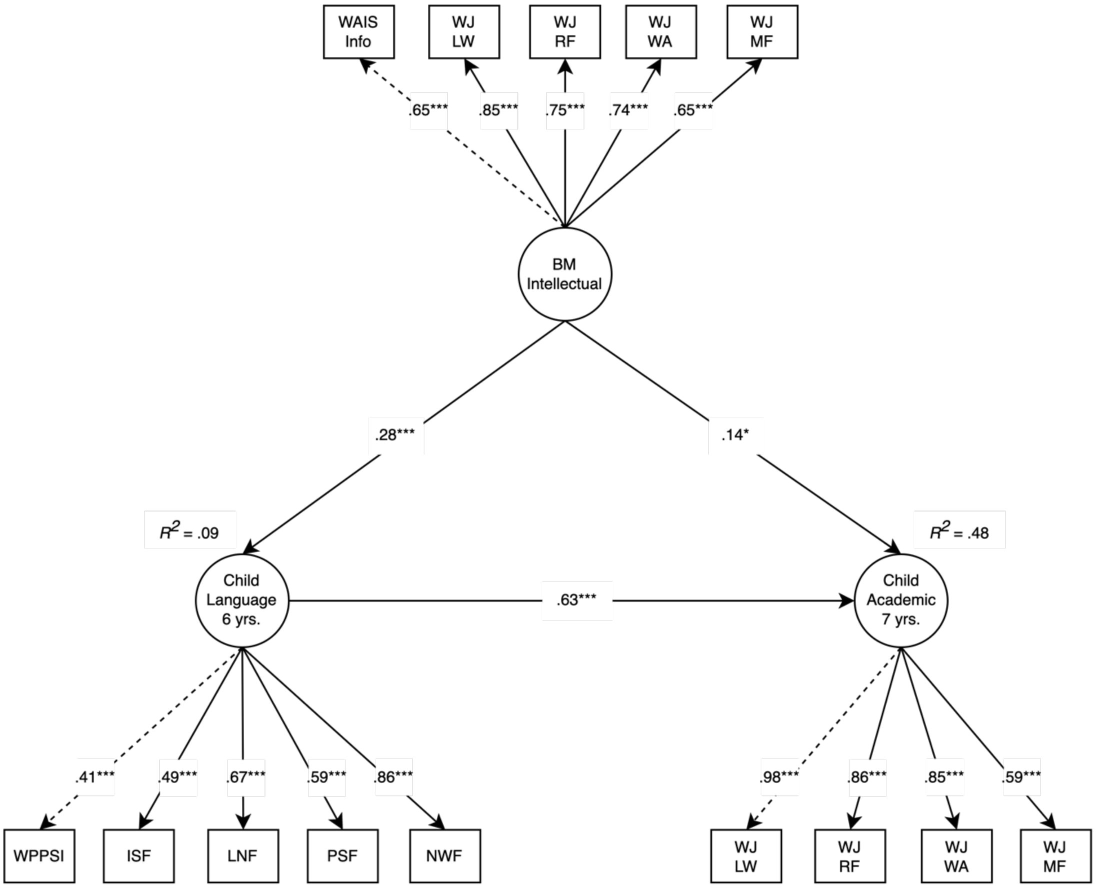
Figure 1B Sensitivity Analysis Re-computing the Birth Mother Analysis Displayed in Figure 2, With (a) Executive Function at 27 Months Dropped from the Model and (b) Executive Function at 27 Months and 54 Months Dropped from the Model



Note. Model fit: (a) $\chi^2(138) = 297.44$, $p < .001$, CFI = .90 RMSEA = .05, SRMR = .06; (b) $\chi^2(81) = 225.96$, $p < .001$, CFI = .91 RMSEA = .06, SRMR = .06. Standardized estimates reported. Faded arrows represent non-significant pathways. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother; EF = executive function; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = maths fluency; FG = forbidden gift; GG = guessing game; DT = dinky toys; G NG = Go NoGo. $^{ns}p \geq .1$ $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

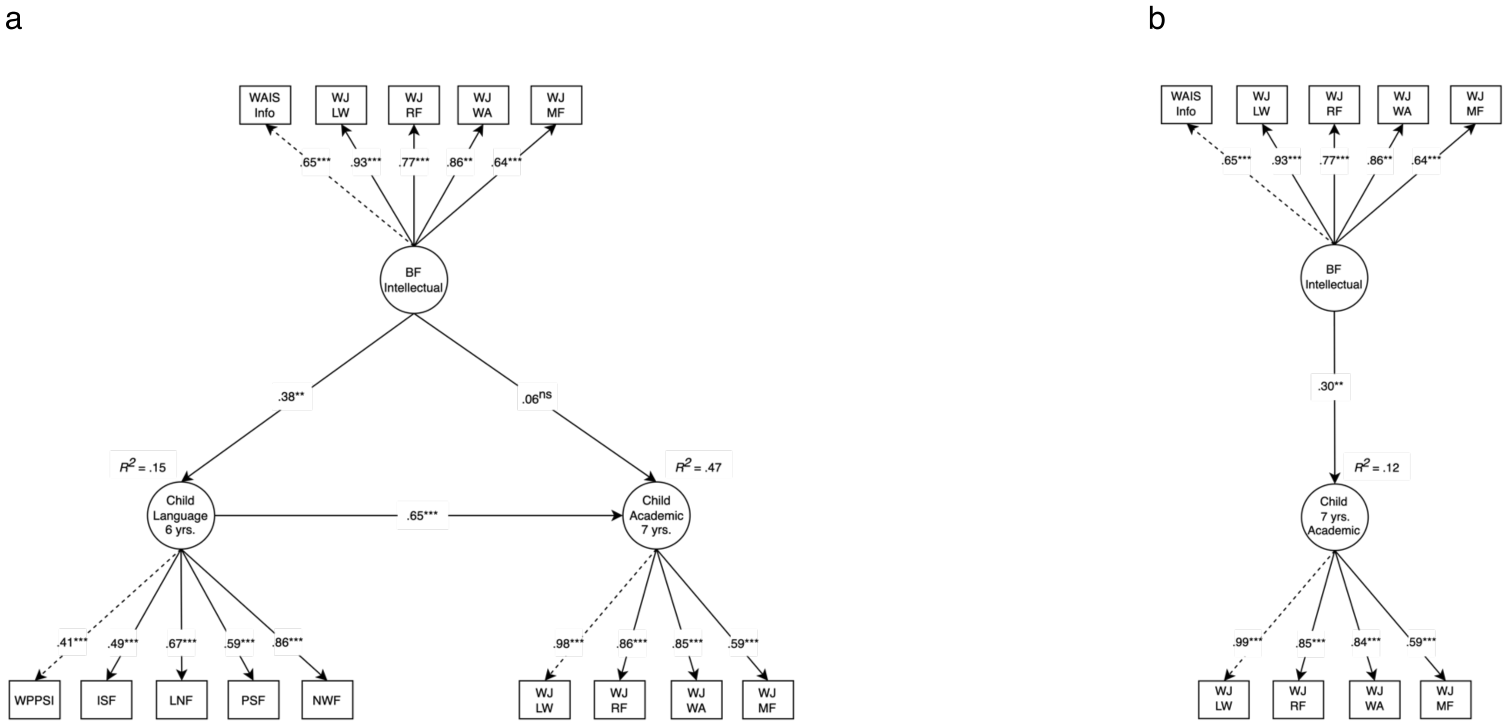
.

Figure 2B Sensitivity Analysis Re-computing the Analysis Displayed in Figure 4, With Language at 27 Months and 4.5 Years Dropped from the Model



Note. Model fit: $\chi^2(110) = 292.22$, $p < .001$, CFI = .90, RMSEA = .06, SRMR = .07. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BM = birth mother; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = maths fluency; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; ISF = Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency; LNF = DIBELS Letter Naming Fluency; PSF = DIBELS Phoneme Segmentation Fluency; NWF = DIBELS Nonsense Word Fluency. $^{ns}p \geq .1$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Figure 3B Sensitivity Analysis Re-computing the Birth Father Analysis Displayed in Figure 5, With (a) Language at 4.5 Years Dropped from the Model and (b) Language at 4.5 and 6 Years Dropped from the Model



Note. Model fit: (a) $\chi^2(110) = 230.88$, $p < .001$, CFI = .92, RMSEA = .05, SRMR = .09; (b) $\chi^2(50) = 101.62$, $p < .001$, CFI = .95, RMSEA = .05, SRMR = .08. Standardized estimates reported. Dashed lines represent parameters that are fixed to 1. Adoption openness, child sex, and obstetric risk were included as covariates in the model. BF = birth father; WAIS Info = Wechsler Adult Intelligence Scale-III Information Subscale; WJ = Woodcock-Johnson III; LW = letter-word association; RF = reading fluency; WA = word-attack; MF = maths fluency; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; ISF = Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Initial Sound Fluency; LNF = DIBELS Letter Naming Fluency; PSF = DIBELS Phoneme Segmentation Fluency; NWF = DIBELS Nonsense Word Fluency. $^{ns}p \geq .1$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Appendix C

C.1 Exploratory Factor Analysis

Prior to hypothesis testing, I examined the factor structure of the parenting measures administered to adoptive parents in EGDS, to determine which items to include in the latent variables constructed for the main analysis. I conducted the exploratory analyses using data from a random split-half subsample (n = 264) at 4.5, 6 and 7 years. The factorability of 39 parenting items was examined. Sixteen of the of the 39 items were from the Alabama Parenting Questionnaire (APQ) (Frick, 1991), 10 of which were from the 'Involvement' subscale of the APQ (measuring parent involvement in the child's daily routine and activities) and 6 of which were from the 'Positive Parenting' subscale of the APQ (measuring the extent to which the parent provides positive feedback or rewards for the child). Six items were from the Warmth subscale of the Iowa Family Interaction Rating Scales (IOWA) (Melby & Conger, 2001), measuring parent warmth towards their child. Eleven of the 39 items were questions on the Home Literacy Environment (HLE) (Johnson et al., 2008; Niklas & Schneider, 2013), measuring family-level HLE characteristics, such as whether the family uses a library card and number of household magazines, as well as child-level HLE characteristics, such as how many books the child owns and how much television they watch. Six items were from the Chaos, Hubbub, and Order Scale (CHAOS) (Matheny et al., 1995), measuring quietness and order of the home.

A correlation matrix of the 39 parenting items revealed that all but 10 items were sufficiently positively or negatively correlated (> .3 or < –.3) with at least one other item at 4.5 and 6 years, and all but 11 items were correlated at 7 years, suggesting reasonable factorability. The items with poor factorability (correlated <.3 or > -.3) were removed from each wave of data. We then used version 2.2.9 of the psych package in R (Revelle, 2022) to conduct Bartlett's Test of Sphericity, which tests whether variables are suitable for data reduction techniques (such as. EFA) by comparing the observed correlation matrix to the identity matrix. The result was significant at each timepoint, suggesting the remaining parenting variables were suitable for EFA: 4.5 years, $\chi^2(406) = 2254$, $p < .001$; 6 years, $\chi^2(406) = 2164$, $p < .001$; 7 years, $\chi^2(378) = 1972$, $p < .001$. To determine how many factors to extract, we created a scree plot at each timepoint of successive eigenvalues from a principal components analysis of the parenting items, using version 2.2.9 of the psych package in R (Revelle, 2022). A widely used decision rule is to retain the factors to the left of the point of inflection in a scree plot of eigenvalues (Cattell, 1966). In the scree plots, the points of inflection were between components 2 and 4, indicating between a 1-factor and 3-factor solution.

Using the R packages psych (version 2.2.9) (Revelle, 2022) and GPArotation (Bernaards & Jennrich, 2005), we ran 3-factor maximum likelihood factor analyses with oblique (oblimin) rotation (as we were expecting the factors to correlate). The results from the three EFAs are displayed in Table 1C. The items that cluster on the same factor suggest that factor 1 represents positive/warm parenting, factor 2 represents screen media use, and factor 3 represents household chaos. Between them, the 3 factors explained 39% of the total variance at 4.5 years, 37% at 6 years, and 38% at 7 years. Factor 1 explained 22% of the variance at each timepoint. Factor 2 explained 9% of the variance at 4.5 and 7 years, and 7% at 6 years. Factor 3 explained 8% of the variance at 4.5 years and 7% at 6 and 7 years. At all timepoints factor 1 was very weakly negatively correlated with factor 2 (4.5 years, $r = -.02$; 6 years, $r = -.01$; 7 years, $r = -.06$) and weakly negatively correlated with factor 3 (4.5 years, $r = -.20$; 6 years, $r = -.23$; 7 years, $r = -.19$), and Factors 2 and 3 were very weakly positively correlated (4.5 years, $r = .04$; 6 years, $r = .07$; 7 years, $r = .13$). All items from the IOWA, and most (but not all) items from the Alabama had high factor loadings on factor 1. Evidence suggests that more indicators per factor is not necessarily better and using too many indicators per factor can create bias in the model chi-square statistic, requiring an increase in sample size to compensate for the increasing model size (Koran, 2020). Consequently, for the main analyses, I decided to drop the items from the APQ and retain the 6 items from the IOWA as indicators in a latent variable measuring positive parenting. The 3 HLE items and single item from the CHAOS that had the highest factor loadings on factor 2 were retained for the latent variable measuring screen media use. The 3 items from the CHAOS that had the highest factor loadings on factor 3 were used as indicators in the latent variable measuring household chaos. Further information on the retained measures is reported in **Chapter 4**.

Table 1C Results from Exploratory Factor Analysis of Parenting Items

	Oblimin Rotated Factor Loadings								
	Positive/Warm Parenting			Screen Media Use			Household Chaos		
Item	4.5 yrs.	6 yrs.	7 yrs.	4.5 yrs.	6 yrs.	7 yrs.	4.5 yrs.	6 yrs.	7 yrs.
You can't hear yourself think in our home.	.01	.02	-.02	-.03	.00	-.06	**.83**	**.84**	**.83**
It's a real zoo in our home.	.02	.04	-.01	.01	-.01	.10	**.90**	**.85**	**.81**
We are usually able to stay on top of things.	.11	.08	.09	-.04	-.17	.02	**-.36**	-.17	-.28
There is usually a television turned on somewhere in our home.	.00	.03	.00	**.58**	**.46**	**.52**	.19	.02	.14
The atmosphere in our house is calm.	.06	.21	.01	-.03	.00	.00	**-.75**	**-.59**	**-.67**
Weekdays (Monday-Friday): On average, how many hours per day does your child watch TV or play video games?	.04	.08	.04	**.69**	**.36**	**.57**	.02	.11	.06
Saturday: On average, how many hours per day does your child watch TV or play video games?	-.03	.01	.01	**.94**	**.88**	**.98**	-.03	.01	.01
Saturday: On average, how many hours per day does your child watch TV or play video games?	.02	-.01	-.03	**.84**	**.92**	**.85**	-.02	-.03	-.04
Let him/her know you really care about him/her	**.57**	**.71**	**.76**	-.03	.01	-.05	-.10	-.05	.14
Act loving and affectionate toward him/her	**.55**	**.70**	**.78**	-.08	.05	-.02	-.06	.04	-.06
Let your child know that you appreciate him/her, his/her ideas, or things he/she does	**.57**	**.82**	**.85**	-.06	-.01	.06	-.05	.02	-.02
Help him/her do something that was important to him/her	**.62**	**.75**	**.75**	-.10	.01	.01	-.02	.04	-.08
Act supportive and understanding toward him/her	**.46**	**.69**	**.63**	.02	.06	.03	-.15	.02	-.08
Tell him/her you love him/her	**.56**	**.65**	**.65**	.02	-.01	.02	.01	.02	.07
You have a friendly talk with your child.	.27	**.47**	**.50**	-.05	.00	.03	-.14	-.07	-.06
You let your child know when he/she is doing a good job with something.	**.65**	**.55**	**.56**	.06	-.09	.07	-.15	.02	-.04
You volunteer to help with special activities that your child is involved in.	**.48**	**.39**	**.35**	-.03	-.10	-.15	.10	-.03	.02
You reward or give something extra to your child for obeying you or behaving well.	**.52**	**.46**	**.41**	-.05	-.12	-.02	-.12	-.09	-.07
You ask your child about his/her day in school.	**.52**	**.49**	**.40**	-.06	.02	-.09	.03	.08	.05
You help your child with his/her homework.	**.42**	**.35**	.20	.10	-.08	.03	.01	-.04	-.01
You compliment your child when he/she does something well.	**.65**	**.49**	**.43**	.10	.09	-.04	-.06	-.20	-.02
You ask your child what his/her plans are for the coming day.	**.48**	**.52**	**.41**	-.07	-.04	-.04	.08	-.07	.18
You drive your child to a special activity.	**.59**	**.39**	**.43**	.01	-.21	-.05	.09	-.03	-.05
You praise your child if he/she behaves well.	**.65**	**.55**	**.48**	.11	.09	.08	.00	-.05	-.06
You hug or kiss your child when he/she has done something well.	**.61**	**.60**	**.63**	.14	.08	-.04	.11	.09	.07
You talk to your child about his/her friends.	**.60**	**.47**	**.54**	-.07	-.05	-.08	.03	-.12	.03
Your child helps plan family activities.	**.41**	**.45**	**.42**	-.07	-.07	-.07	.08	.03	.03
You attend PTA meetings, parent/teacher conferences, or other meetings at your child's school.	**.50**	.21	.11	-.15	-.12	-.11	.09	-.13	-.10
You tell your child that you like it when he/she helps out around the house.	**.59**	**.47**	NA	.01	-.15	NA	.09	-.04	NA
Eigenvalues	6.25	6.52	6.15	2.53	2.15	4.40	2.41	2.00	2.06
Proportion of variance	.22	.22	.22	.09	.07	.09	.08	.07	.07

Note. Factor loadings over .40 appear in bold