Socioeconomic factors and common mental health disorders: The role of gene-environment interplay

Badini, Isabella

Awarding institution:
King's College London

Socioeconomic factors and common mental health disorders: The role of gene-environment interplay

Isabella Badini

Social, Genetic and Developmental Psychiatry Centre,
Institute of Psychiatry, Psychology and Neuroscience

PhD thesis, 2023

Submitted for the degree of Doctor of Philosophy in
Social, Genetic and Developmental Psychiatry,
King's College London

Abstract

This thesis aims to enhance our understanding of the mechanisms underlying socioeconomic disparities in common mental health outcomes by investigating gene-environment interactions and the impact of selection bias in family data. Significant gaps exist in the literature regarding the moderating effects of socioeconomic status on the aetiology of mental health outcomes and the potential genetic overlap between family socioeconomic conditions and child mental health. Moreover, limited research has explored gene-environment interplay processes in lower- and middle-income populations. Additionally, selection bias in population-based cohort studies can compromise the validity of study results and potentially limit our understanding of the effects of socioeconomic conditions on mental health.

Chapter 2 uses extended family data from the Norwegian Mother, Father and Child Birth Cohort Study (MoBa) to examine how parental socioeconomic factors moderate the aetiological influences on child emotional and behavioural problems in the presence of gene-environment correlation. Chapter 3 investigates the influence of individual socioeconomic factors on depression symptoms in adults using data from the Colombo Twins and Singletons Study (CoTASS) in Sri Lanka, providing insights into the relationship between socioeconomic status and mental health outcomes in a South Asian context. Chapter 4 explores the impact of selection bias on phenotypic and genetic correlations within the MoBa cohort, assessing the potential biases introduced by selective participation.

By addressing these gaps in the literature and considering the role of gene-environment interactions and selection bias, this thesis aims to advance our knowledge of the complex relationships between socioeconomic conditions and mental health outcomes. The findings have implications for interventions and policies targeting mental health disparities and promoting well-being across diverse populations.

Table of Contents

Table of Figures

Chapter 2:

Chapter 3:

Chapter 4:

Table of Tables

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Acknowledgements

The research presented in this thesis was made possible by the thousands of people who volunteered their time to take part in research, as well as the research teams who conducted, managed, built and maintained each cohort. This work is thanks to the time and commitment invested by all those involved and to the funders of each study. My position as a PhD student at King's College London was funded by the Economic and Social Research Council (ESRC) (grant number: ST11872).

I owe immeasurable gratitude to my supervisors, Dr. Tom McAdams and Dr. Helena Zavos, for their invaluable contributions that have been instrumental in the successful completion of my doctoral research. I am truly thankful for their unwavering support, guidance, and expertise throughout this journey. The time they devoted to mentoring me and their provision of exceptional training and funding have been of immense significance. Their collective efforts have profoundly shaped my academic growth and instilled in me the confidence to overcome challenges, for which I am grateful for.

I am incredibly fortunate to have benefitted from the support of all department leads, staff, researchers and students at The Social, Genetic and Developmental Psychiatry Centre, who have built and maintained an open, engaging and supportive community. I recognise and appreciate the wealth of knowledge and experiences I have gained through my interactions with my colleagues and peers over the years. Special thanks to my team members and peers Yasmin Ahmadzadeh, Chris Rayner, Joanna Bright, Daniel Wechsler, Anna Furtjes, and Millie Ward, who have taken the time to share their knowledge, ideas and skills with me.

The mentorship that I have received for each of my PhD projects has also extended from individuals across countries and institutions. I am immensely grateful for the generosity and support of Eivind Ystrom and the IToR team at the University of Oslo, Norway, and the CoTASS research team at the Insititute for Research and Development in Health and Social Care, Colombo, Sri Lanka.

Finally, thank you to Lucia, Gabriele, Olivia, i nonni, and close friends, for their unwavering, love, support and encouragement throughout my PhD journey.

Statement of Authorship

All work presented in this thesis is my own except where acknowledged in the text. All investigations were carried out by me, as first author, in collaboration with colleagues included in the author lists at the start of each chapter. Chapter-specific author contributions are described below. Data collection for the samples used across all chapters were completed by the respective research teams prior to my involvement.

Isabella Badini

Chapter 2:

IB and TM conceived and designed the investigation. TM and EME developed models for use in the analyses. IB carried out statistical analyses with support from TM. IB drafted and revised the manuscript, and all co-authors critically reviewed the manuscript.

Chapter 3:

IB, HZ and FVR conceived and designed the investigation. FVR and HZ developed models for use in the analyses. IB carried out statistical analyses with support from HZ. IB drafted and IB and HZ revised the manuscript, and all authors critically reviewed the manuscript.

Chapter 4:

IB and TM conceived and designed the investigation. IB carried out statistical analyses with support from TM and HZ. IB wrote and revised the manuscript, and IB, TM and HZ discussed results and critically reviewed the manuscript.

Publications arising from the chapters in this thesis

Chapters 2 and 3 have undergone peer-review and have been accepted for publication:

Badini, I., Ahmadzadeh, Y. I., Wechsler, D. L., Lyngstad, T. H., Rayner, C., Eilersten, E. M., Zavos, H. M.S., Ystrom, E., McAdams, T. A. (*in press*). Socioeconomic status and risk for child psychopathology: Exploring gene-environment interaction in the presence of gene-environment correlation using extended families in the Norwegian Mother, Father and Child Birth Cohort Study. *Journal of Child Psychology and Psychiatry.*

Badini, I., Jayaweera, K., Pannala, G., Adikari, A., Siribaddana, S., Sumathipala, A., McAdams, T. A., Harber-Aschan, L., Hotopf, M., Rijsdijk, F. V., & Zavos, H. M. S. (2023). Associations between socioeconomic factors and depression in Sri Lanka: The role of gene-environment interplay. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2023.07.084

Other publications

Zavos, H. M. S., Zunszain, P. A., Jayaweera, K., Powell, T. R., Chatzivasileiadou, M., Harber-Aschan, L., Adikari, A., Pannala, G., Siribaddana, S., **Badini, I.,** Pariante, C. M., Hotopf, M., Sumathipala, A., & Rijsdijk, F. (2022). Relationship between CRP and depression: A genetically sensitive study in Sri Lanka. *Journal of Affective Disorders*, *297*, 112–117. https://doi.org/10.1016/j.jad.2021.10.003

Wechsler, D. L., Rijsdijk, F. V., Adamo, N., Eilertsen, E. M., Ahmadzadeh, Y. I., **Badini, I.,** Hannigan, L. J., Ystrom, E., & McAdams, T. A. (2022). *Evidence for Strong Aetiological Overlap Between Child and Adult ADHD Symptomatology Using an Extended Family Design* (SSRN Scholarly Paper No. 4017374). https://doi.org/10.2139/ssrn.4017374

Badini, I., Coleman, J. R. I., Hagenaars, S. P., Hotopf, M., Breen, G., Lewis, C. M., & Fabbri, C. (2020). Depression with atypical neurovegetative symptoms shares genetic predisposition with immuno-metabolic traits and alcohol consumption. *Psychological Medicine*, *52*(4), 726–736. https://doi.org/10.1017/s0033291720002342

1. Background

1.1. Introduction

Mental health disorders pose a significant global burden, affecting individuals across the lifespan and exerting wide-ranging impacts on societies and individuals. Mental health disorders contribute to approximately 10% of the global disease burden, with depression and anxiety being major causes of disability worldwide (WHO, 2022). These conditions typically emerge during childhood or adolescence and, if untreated, can have adverse effects on psychosocial functioning and healthcare demands. This includes the impact of mental health disorders on individuals' overall health, functioning, and quality of life. The global burden of mental health disorders also includes societal and economic costs associated with reduced productivity, increased healthcare utilisation, and social consequences (Arias et al., 2022; GBD 2019 Mental Disorders Collaborators, 2022). Mental health disorders can lead to significant disability, impairment in daily functioning, and increased mortality rates in some cases (GBD 2019 Mental Disorders Collaborators, 2022; WHO, 2022). Co-occurring psychiatric disorders are common, often leading to greater symptom severity and poorer outcomes (Daré et al., 2019; McGrath et al., 2020; Plana-Ripoll et al., 2019). Addressing the burden of mental health disorders requires a comprehensive approach that includes promoting mental health awareness, improving access to quality mental health services, reducing stigma, and implementing evidence-based interventions and policies (WHO, 2022). Understanding the causes of mental health disorders is crucial in addressing their burden effectively (Alegría et al., 2018; Andreassen et al., 2023; Maselko, 2017). Identifying and understanding the underlying causes and risk factors can aid in the development of targeted interventions, prevention strategies, and treatment approaches to mitigate the impact of mental health disorders.

In this thesis, I aim to explore the relationship between socioeconomic factors and common mental health outcomes using genetically informed research methods. Specifically, the studies included in this thesis focus on understanding the mechanisms underpinning the observed relationships between family socioeconomic conditions and internalising and externalising

problems in children, and internalising disorders in adults. In this first chapter, I provide a brief overview of the prevalence, characteristics and aetiology of internalising and externalising symptoms. I then discuss common indicators of socioeconomic status and summarise findings from epidemiological research investigating the link between socioeconomic conditions and mental health outcomes in children and adults. I also provide an overview of proposed theories and mechanisms underlying socioeconomic disparities in mental health. To end, I outline three main challenges addressed in this thesis for advancing research efforts and provide a summary of the thesis aims and structure.

1.2. An introduction to internalising and externalising problems

1.2.1. Prevalence and characteristics of internalising problems in adults

Internalising disorders, such as anxiety and depression, constitute the most common mental health conditions worldwide (Kessler et al., 2009; World Health Organization [WHO], 2017). Prevalence estimates vary across studies and populations, but generally, anxiety disorders has been estimated to affect around 4% of adults globally, while the global prevalence of depressive disorders is estimated to be 3.8% (WHO, 2022). Characteristics of internalising disorders in adults include persistent feelings of sadness, hopelessness, worry, and loss of interest or pleasure in activities (American Psychiatric Association [APA], 2013; WHO, 2015). Anxiety disorders typically become prevalent at an earlier age than depressive disorders, and both become increasingly common in later life (Kalin, 2020; Kessler et al., 2007). Internalising symptoms exhibit high comorbidity across diagnostic groups, with around half of individuals with depressive disorder also diagnosed with an anxiety disorder (Kalin, 2020; Kessler et al., 1998; Kroenke et al., 2007).

1.2.2. Prevalence and characteristics of internalising and externalising problems in children

In children, internalising problems manifest as emotional difficulties, such as excessive shyness, withdrawal, and somatic complaints (Achenbach & Edelbrock, 1983; APA, 2013). Anxiety and depression are among the most common childhood psychiatric disorders, with global prevalence estimates ranging from 1.3% to 3.2% across ages 5-19 years (WHO, 2022).

Externalising problems, such as conduct disorders and attention-deficit/hyper-active disorder (ADHD), are also prevalent among children worldwide. Research conducted in 2019 indicated that ADHD affected approximately 0.3 to 2.6% of children and adolescents, while the global prevalence of conduct disorders was estimated to be around 2.3% (WHO, 2022). Externalising problems typically manifest as impulsive and disruptive behaviours, aggression, difficulties in self-control, and poor attention span (Achenbach & Edelbrock, 1983; APA, 2013). Childhood onset internalising and externalising problems have been associated with impairments in social and academic functioning, and with later mental health disorders and worse prognoses in adulthood (Jaffee et al., 2002; Last et al., 1997; Moffitt & Caspi, 2001; Pedersen et al., 2019; Pine et al., 1998).

1.2.3. Aetiology of internalising and externalising problems

Both genetic and environmental factors contribute to the aetiology of internalising and externalising problems (Q. Chen et al., 2017; Hettema et al., 2001; Kendall et al., 2021; Polderman et al., 2015). Family, twin and adoption studies have provided evidence for the heritability of these disorders (Knopik et al., 2016; Plomin et al., 2016; Polderman et al., 2015). Family-based genetically informative designs leverage knowledge of relatedness among individuals within a family (e.g., adopted children, siblings, twins, and parents) to make inferences about the contribution of genetic and environmental factors underlying observed phenotypic variance in a trait and/or associations between traits. Heritability is defined as the proportion of variation in a trait that is attributable to genetic differences among individuals in a specific sample drawn from the larger population at a particular time. Twin studies have predominantly been used to derive trait heritability within a population. The twin design involves comparisons of reared-together identical (monozygotic; MZ) twins - who share virtually 100% of their genetic material, and non-identical (dizygotic; DZ) twins - who share ~50% of their segregating genes (i.e., genetic material that differs between individuals). Holding environmental influences constant (i.e., assuming that MZ and DZ twin pairs experience their shared environment equally), the phenotypic variance and/or covariance between traits can be decomposed into genetic and environmental components (Rijsdijk & Sham, 2002).

Twin studies have provided important implications for our understanding of the aetiology of psychiatric traits. Heritability estimates reported in twin studies range from around 30-50% for anxiety and depression, to over 80% for ADHD (Q. Chen et al., 2017; Hettema et al., 2001; Kendall et al., 2021; Polderman et al., 2015). Further, these findings provide the strongest evidence for the role of non-genetic influences on mental health after accounting for genetic factors. Environmental factors, including stressful life events, socioeconomic factors, cultural influences, and adverse childhood experiences, interact with genetic predispositions to influence the development of psychiatric disorders (Assary et al., 2018; Dick, 2011; Plomin et al., 2016; Rutter et al., 2006; Shanahan & Hofer, 2005). Notably, socioeconomic status is considered a key social determinant of mental health, influencing the development, prevalence and severity of internalising and externalising disorders (Braveman & Gottlieb, 2014; Marmot & Equity, 2014; Maselko, 2017; Wilkinson et al., 2003). In the following section of this chapter, I provide an overview of common indicators of socioeconomic status and summarise study findings on the link between socioeconomic factors and mental health.

1.3. Socioeconomic status and mental health

1.3.1. Defining and measuring socioeconomic status

Socioeconomic status (SES) is considered a key explanatory factor across disciplines such as social science, health sciences, and education research, due to its associations with physical and mental health outcomes, as well as cognitive and socio-emotional domains (Alegría et al., 2018; Eriksson et al., 2021, 2021; Flaskerud & DeLilly, 2012; Morgan et al., 2009; Phelan et al., 2010). It reflects individual and/or group (e.g., household) access to and control over (*normatively valued*[1]) social and economic resources (Antonoplis, 2023; Diemer et al., 2013). SES is conceptualised through indicators or measures collected at the individual, household or area-based (e.g., neighbourhood) level (Conway et al., 2019; Cooper et al., 2012; Diemer et al., 2013; Howe et al., 2012). Researchers use these indicators to capture and analyse the effects of socioeconomic factors on psychological and life outcomes. The appropriate measurement of SES and its association with mental health outcomes has been shown to vary

[1] In this context, *normatively valued* means that certain aspects of socioeconomic status (SES) are considered important and desirable within a specific societal framework (Antonoplis, 2023). It refers to the features of SES that are culturally and socially recognised as beneficial for individuals to effectively participate in and conform to the expectations and norms of their particular time and place.

according to the specific research question, study sample and context (Cooper et al., 2012; Diemer et al., 2013; Festin et al., 2017; Maselko, 2017; F. Reiss, 2013). Several reviews of SES indicators have discussed the strengths and limitations of different measurement approaches (Antonoplis, 2023; Cooper et al., 2012; Diemer et al., 2013; Howe et al., 2012; Maselko, 2017). Drawing from these reviews, I provide an overview of commonly used indicators in the measurement of SES, including examples of their interpretation. For the purpose of this thesis, I concentrate on measures of individual or household-level SES rather than area-based (e.g., neighbourhood disadvantage) measures.

Traditional SES measurement

When studying the effects of socioeconomic conditions on mental health outcomes, researchers use a range of indicators intended to capture access to and control over social and material resources (Diemer et al., 2013). Indicators of SES generally cluster around two main domains: (1) prestige, which reflects social standing and occupational prestige; and (2) resources, which encompass material (i.e., physical possessions or properties that have value) and financial (i.e., monetary resources or investments) assets. Typically, these indicators include one or a combination of individuals' occupational status (prestige), income and education (resources) (Diemer et al., 2013). Table 1 provides an overview of these SES indicators, including examples of common measurement approaches and notes on their interpretation.

Measuring SES in low- and middle-income countries

Due to differences in social stratification processes, cultural factors and data availability, the appropriate measurement of SES varies across high-income and low- and middle-income countries (LMIC) (Cooper et al., 2012; Howe et al., 2012). Conventional measures of SES in higher income populations are often problematic for analyses in lower- and middle-income populations (Cooper et al., 2012; Howe et al., 2012). For example, income and occupation can be more difficult to measure in many LMIC due to the presence of a larger informal economy, casual and seasonal labour, and wide-spread rural dwelling (Cooper et al., 2012). Instead, measures of household assets, education, financial strain, and food insecurity are commonly used indicators of SES in lower- and middle-income contexts (Cooper et al., 2012; Howe et al., 2012; Maselko, 2017). Research in these settings has shown that studies using domains such

as educational attainment or household assets yield more consistent results than those using income (Araya et al., 2003; Howe et al., 2012; Maselko et al., 2018). Table 2 provides an overview of these SES indicators, including examples of common measurement approaches and notes on their interpretation.

Table 1. Individual indicators of socioeconomic status commonly used in high income countries

Indicator	Measurement methods	Interpretation
Education	Educational attainment: individual or household level; highest level attained; qualifications; years completed Continuous or categorical International Standard Classification of Education (ISCED)	Reflects early-life SES, usually stable across the life-course Strong determinant of employment and income Reflects material, cultural, and other resources of the family of origin Captures long-term influences of both early life circumstances on adult outcomes, as well as influences of adult resources (e.g., through employment status or income) on outcomes Affects access to health care or information Comparable data across multiple countries Context-specific: country education system, age cohorts
Income	Sum of income: individual or household level; monthly or annual sum; before taxes Sources of income typically include earnings, government or state welfare benefit support, pensions and interest Absolute or relative poverty thresholds	Measures access to material resources (e.g., food, shelter, and culture) and access to services (e.g., health care, leisure, and education) Relates to social standing/prestige Context-specific: country, age, sex
Occupation	Employment or job history: individual or household level – either highest household occupation or averaged; current or most recent; Classified into occupational categories, according to resources such as Census classification systems; the National Statistics Socio-economic classification (NS-SEC); Standard Occupational Classification (SOC)	Reflects social standing or prestige, working relations and conditions Strong determinant of income Based on educational attainment and social resources (e.g., social network and social ties) Influences social networks, work-based stress, autonomy or control Excludes some groups e.g., retired people, unpaid home workers or care givers, students Context-specific: country (e.g., level of industrialisation), age cohorts

Note. Key papers summarising best practices in conceptualising and measuring socioeconomic status include: Diemer et al. (2013) and Antonoplis (2023)
Source: Adapted from Diemer et al. (2013)

Table 2. Individual indicators of socioeconomic status commonly used in low- and middle-income countries

Indicator	Measurement methods	Interpretation
Education	Educational attainment: individual or household level; highest level attained; qualifications; years completed Continuous or categorical International Standard Classification of Education (ISCED)	Reflects early-life SES, usually stable across the life-course Strong determinant of employment and income Reflects material, cultural, and other resources of the family of origin Captures long-term influences of both early life circumstances on adult outcomes, as well as influences of adult resources (e.g., through employment status or income) on outcomes Affects access to health care or information Comparable data across multiple countries Context-specific: country education system, age cohorts
Asset-based measures	Ownership of assets: household level; durable assets (e.g., car, refrigerator, television), housing characteristics (e.g., tenure of dwelling, main material of floor and roof, main cooking fuel), and access to basic services (e.g., main electricity supply, source of drinking water, sanitation facilities) Demographic and Health Surveys (DHS) Asset index: sum or weighted sum of items included	Captures material aspects of living conditions Interpretation depends on the relationship of individual to the household (e.g., family or parents' SES for children and young adults still in the family home, or spousal household SES for married women not in paid employment and living in spouse's family dwelling) Reflects financial standing and resources available to the individual Relatively stable measure of SES, varies less in response to fluctuations in income and expenditure and resistant to many economic shocks Comparable data across multiple countries Country-specific: asset items included, age cohorts
Financial stress	Self-reported stress due to financial difficulties; example of survey question and response options: *"How well do you feel you are managing financially these days?* *1… Living comfortably* *2… Doing alright* *3… Just about getting by* *4… Finding it difficult to make ends meet* *5… Finding it very difficult to make ends meet"* Food insecurity Individual or household debt	Captures self-reported stress, worry or anxiety related to financial difficulties or hardships Reflects economic challenges individuals face in meeting their basic needs and managing financial responsibilities Availability, access and affordability of food Indicator of resource deficiency and negatively associated with subjective wellbeing

Note. Key papers summarising best practices in conceptualising and measuring socioeconomic status in low- and middle-income countries include: Howe et al. (2013) and Cooper et al. (2012)

1.3.2. Overview of epidemiological studies on associations between socioeconomic factors and mental health outcomes in children and adults

A substantial body of research has demonstrated the impact of childhood socioeconomic disadvantage on mental health (Costello et al., 2003; Fitzsimons et al., 2017; Kinge et al., 2021; Lansford et al., 2019; McLaughlin et al., 2011; Melchior et al., 2014; Peverill et al., 2021; Pryor et al., 2019; F. Reiss, 2013; F. Reiss et al., 2019). A systematic review from 23 countries found that children from families with low SES had a two- to three-fold increased risk of developing emotional and behavioural disorders compared to children from more advantaged families (F. Reiss, 2013). Experience of socioeconomic disadvantage in childhood has also been associated with negative outcomes for physical health, cognitive development, educational attainment, and social well-being (Clarke et al., 2022; Currie, 2009; Duncan et al., 1998; Duncan & Magnuson, 2003, 2003; Lee & Jackson, 2017). These associations have been found to start in early childhood and continue into adulthood (Hakulinen et al., 2020; McLaughlin et al., 2011; Melchior et al., 2014; Miller et al., 2021; Najman et al., 2004, 2010; Whitfield et al., 2021).

When studying the effects of family socioeconomic conditions on child development, researchers have used various indicators to assess the family's access to material and social resources (Diemer et al., 2013). These typically include one or a combination of parental income, educational attainment, and occupational status (Diemer et al., 2013; Duncan & Magnuson, 2003). While these indicators are moderately correlated, each has been shown to capture distinctive aspects of the socioeconomic environment (Duncan & Magnuson, 2003; Geyer et al., 2006). For example, family income is often used as a proxy measure for material hardship, and parental education provides a measure of human capital in the family that is typically more stable than other components of SES (Diemer et al., 2013; Peverill et al., 2021). Thus, different indicators of SES may share some common associations with mental health outcomes while also having distinct effects. Past studies comparing socioeconomic indicators have shown that parental income and educational attainment have stronger associations with child mental health outcomes than other measures of family SES (Lansford et al., 2019; McLaughlin et al., 2012; F. Reiss, 2013). In a longitudinal study examining childhood socioeconomic factors and the development of mental health problems, low family income emerged as a significant predictor of disorder onset, while low parental education was

associated with greater persistence and severity of mental health difficulties (McLaughlin et al., 2012). Supporting these findings, a prospective register-based study in Norway, comprising children aged 5-17 years, showed that the association between parental income and offspring mental health diagnoses persisted even after considering parental education (Kinge et al., 2021).

Possible explanations for these findings are that the effect of household income translates into material and immaterial resources that promote healthy child development, such as adequate housing conditions, food, healthcare, resources for managing stressful and demanding circumstances (e.g., by seeking professional help), and resources to invest in parent-child interactions (Duncan & Magnuson, 2003; Festin et al., 2017; Geyer et al., 2006). Conversely, parental education can be taken as an indicator of childhood social environment, as parent's educational background may influence parenting behaviours, parental investment, availability of learning materials in the home, and health literacy (e.g., access and understanding of therapeutic measures) - all of which may influence risk for psychopathology (Duncan & Magnuson, 2003; Festin et al., 2017; Geyer et al., 2006).

Prior evidence suggests that the strength of association between family SES and child psychopathology varies depending on the type of disorder, with externalising disorders showing a stronger link with low SES compared to internalising disorders (F. Reiss, 2013). However, the extent of this difference varies across different study populations (F. Reiss, 2013), and some population-representative studies have found little evidence of moderation (Miller et al., 2021) or have shown that family socioeconomic factors have stronger associations with internalising symptoms (Vollebergh et al., 2006). This suggests that differences observed in study results may be accounted for by other (interacting) factors, such as the age of the sample or the specific measure used to assess family SES (Peverill et al., 2021; F. Reiss, 2013). For example, studies have found that family socioeconomic factors were associated with internalising and externalising symptoms in early childhood, but that associations with internalising symptoms were attenuated with increasing age (Houweling et al., 2022; Strohschein, 2005).

Exposure to childhood socioeconomic disadvantage has been associated with increased risk for mental health problems later in life in longitudinal studies (Hakulinen et al., 2020; McLaughlin et al., 2011; Melchior et al., 2014; Miller et al., 2021; Najman et al., 2010; Whitfield et al., 2021). For example, household income volatility in childhood has been related to increased externalising problems in adolescence (Miller et al., 2021; Whitfield et al., 2021). A recent Danish national cohort study conducted a longitudinal analysis to examine the relationship between parental income and income mobility in childhood, and subsequent risk of developing mental health disorders in adulthood (Hakulinen et al., 2020). The findings showed that around a quarter of individuals from the lowest income families were diagnosed with a mental health disorder by age 37, compared with 13% of those born into the highest income families. Higher risk for developing a mental health disorder in adulthood was associated with more time spent in economically disadvantaged circumstances and downward family income mobility during childhood.

Epidemiological research in adults has consistently shown higher levels of mental health problems in individuals with greater socioeconomic disadvantage (Burns, 2015; Domènech-Abella et al., 2018; Kivimäki et al., 2020; Lund, 2014; Lund et al., 2010; Maselko, 2017; Patel et al., 2018). Whilst evidence of a relationship is broadly supported, the strength of associations between different socioeconomic indicators and mental health outcomes varies across studies and settings (Alegría et al., 2018; Darin-Mattsson et al., 2017; Diemer et al., 2013; Korous et al., 2022; Maselko, 2017; Maselko et al., 2018). Methodological characteristics such as the use of different socioeconomic indicators (or composite SES indices), measurement error, and characteristics of the study sample are likely to contribute to variation in the magnitude of the observed associations (Angel et al., 2019; Festin et al., 2017; Hoebel et al., 2017). Research has shown only moderate correlations between different socioeconomic indicators, and there is evidence from studies conducted in high-income populations that using domains such as income and educational attainment yield more consistent results than other measures of SES (Araya et al., 2003; Howe et al., 2012). In contrast, in low-middle income countries (LMIC), data on factors such as income are often unreliable, and indicators such as household assets, educational attainment and financial stress are seen as better indicators due to different social stratification processes (Cooper et al., 2012; Howe et al., 2012; Lund et al., 2010; Maselko et al., 2018). For example, a study in

rural Pakistan found that fewer assets, food insecurity, and household debt were associated with more severe prenatal depression symptoms (Maselko et al., 2018). As each of these indicators capture a different aspect of SES, with a potentially different pathway linking it to mental health, heterogenous findings both within and between countries could be expected.

1.4. Mechanisms underlying the relationship between socioeconomic status and mental health

1.4.1. Social selection and social causation

Two principal pathways are thought to underlie the observed associations between SES and mental health outcomes: social causation and social selection (Dohrenwend et al., 1992). The social causation hypothesis posits that exposure to the adverse social and economic conditions associated with lower SES (such as poor environmental conditions, material and social deprivation, and increased exposure to adverse and stressful life events) increases risk for mental health conditions (Dohrenwend et al., 1992). The social selection hypothesis instead suggests that individuals with mental health problems are more likely to drift into or remain in lower SES levels due to disability, reduced economic productivity, loss of employment, increased health expenditure, and stigma as a result of their illness.

Longitudinal analyses have been applied to investigate the nature of the relationship between socioeconomic status and mental health. For example, a recent longitudinal study investigating the directionality of the relationship between SES and depression in a nationally representative sample from South Africa found evidence for a reciprocal relationship between SES and depression in LMIC (Lund & Cois, 2018). Socioeconomic disadvantage at baseline was associated with worse depression symptoms at two- and four-year follow-up assessments, and worse depression symptoms at baseline was associated with lower SES two years later. The findings from this study suggest that both social causation and social selection mechanisms act simultaneously to reinforce cycles of poverty and depression.

Poverty alleviation programs such as cash transfer interventions have provided evidence on whether improvements in socioeconomic conditions can lead to improved mental health outcomes. These interventions address the social causation pathway. For example, the

Oportunidades program was an intervention in Mexico where women received a stipend that constituted roughly 25% of household income (Ozer et al., 2011). At the time of evaluation, women had been in the program from three to five years. Depression symptoms were significantly reduced in the intervention group and this effect was partially mediated through a reduction in stress and increase in perceived control. Furthermore, experimental evidence has implicated social causationas a mechanism underpinning the SES-mental health relationship in children. In a natural experiment where Native American families received income supplement following the opening of a casino, there was a significant reduction in children's behavioural (externalising) symptoms (Costello et al., 2003) and emotional (internalising) symptoms (Akee et al., 2018). These findings suggest causal influences are possible because the income intervention that moved some families out of poverty cannot be ascribed to unobserved familial confounders (i.e., genetic and environmental factors shared by family members).

There is clear evidence from existing correlational research that socioeconomic conditions predict mental health outcomes, whereby those at the lower end of the socioeconomic distribution experience worse health than those at the higher end. Further, findings from longitudinal and experimental research are suggestive of causal influences of socioeconomic factors on mental health. However, genetic factors influence nearly all human traits and experiences to some degree (Plomin et al., 2016). A key limitation of the literature is that analysis of the relationship between SES and mental health can be influenced by unobserved familial confounders (i.e., genetic and environmental factors shared by family members) in studies of unrelated individuals. To better understand the mechanisms underlying this relationship, it is important to consider the interplay between genetic and environmental factors.

1.4.2. Gene-environment interplay

In the following section, I provide an overview of gene-environment interplay, which includes gene-environment interaction (GxE) and gene-environment correlation (rGE) (Dick, 2011; Rutter, 2006). GxE refers to a situation in which aetiological influences on a trait that are moderated by contextual factors, while rGE refers to the influence of an individual's genotype on their exposure to environmental factors. These concepts are essential for understanding

how socioeconomic factors interact with genetic (and environmental) influences to impact mental health outcomes. Additionally, I discuss theoretical frameworks that shed light on the nature of GxE, such as the diathesis-stress model, differential susceptibility model, bioecological (social compensation) model, and social control/push model.

1.4.2.1. Gene-environment correlation (rGE): genetic influences associated with exposure to (socioeconomic) context-environment

Observed associations between socioeconomic factors and mental health outcomes have often been interpreted as direct environmental influences. However, associations between socioeconomic factors and mental health problems could also arise because they share common causes. Genetic influences have been shown to be associated with a variety of environmental exposures (Kendler and Baker, 2007). This is known as gene-environment correlation (rGE) and describes genetically influenced behaviour which can influence individuals' exposure to certain environments.

Genetically-informative research has shown moderate genetic influences on indicators of SES (Ball et al., 2010; Hill et al., 2019; Marees et al., 2021; Rimfeld et al., 2018). For example, twin studies have shown that genetic factors explain ~40% of the phenotypic variance in educational attainment, and 39-54% of the variance in lifetime earnings (Hyytinen et al., 2019; Silventoinen et al., 2020). Consistent evidence from genome-wide association studies has shown that common genetic variants explain 11% and 15% of the variance in household income and educational attainment, respectively (Hill et al., 2019). If there is an overlap in the genetic factors associated with mental health problems, then part of the link between SES and mental health could be explained by common causes. Genetic correlations between SES indicators and various psychiatric disorders have been reported based on genomic data (e.g., income and depression [rg = -0.24]; income and ADHD [rg = -0.47]; education and depression [rg = -0.08]; education and ADHD [rg = -0.52]; Hill et al., 2019). These findings could be taken as support for the social selection hypothesis, in which affected individuals may be more likely to drift into or remain at lower SES levels, at least in part, based on genetically influenced traits and behaviours related to mental health outcomes.

Familial confounding can introduce bias into estimates and result in incorrect conclusions regarding the causal relationship between family socioeconomic factors and childhood psychopathology. Studies have shown that child emotional and behavioural problems (Burt, 2009; Nikstat & Riemann, 2020; Rice et al., 2002) and family socioeconomic indicators (Krapohl & Plomin, 2016; Trzaskowski et al., 2014) are heritable, and that at least part of the association between them can be attributed to common genetic influence (Hill et al., 2019; Krapohl et al., 2017; Torvik et al., 2020; Visscher et al., 2017). Because parents pass on genotypes to their biological children, as well as provide family environments that correlate with their genotypes, the association between family SES and children's mental health outcomes can be genetically mediated (i.e., passive gene-environment correlation; Kendler & Baker, 2007; Plomin et al., 1977).

Behavioural genetic designs such as sibling, adoption, and children-of-twins (CoT) designs can distinguish between shared familial (i.e., genes and environments shared between parent and child) and phenotypic mechanisms of intergenerational association (McAdams et al., 2014; McAdams et al., 2018). A recent study using an extension of the CoT design showed that intergenerational associations between parental education and depression could be fully accounted for by genetic factors shared between parents and their children (Torvik et al., 2020). Shared familial factors were also found to contribute to associations between parental education and child ADHD.

1.4.2.2. Gene-environment interaction (GxE): genetic (and environmental) influences moderated by the socio-environmental context

Socioeconomic conditions may also affect the relative importance of genetic and environmental influences on mental health within a population (Rutter, 2006). This is known as gene-environment interaction (GxE) and could reflect a form of social causation whereby aetiological influences on a trait are moderated by environmental context (Dick, 2011; Rutter, 2006).

Early attempts to model GxE using twin data would stratify the sample by level of the environmental variable and estimate genetic and environmental influences as a function of these subgroups (Neale & Cardon, 1992). However, there are several problems with this

approach (Dick, 2011; Purcell, 2002). First, modelling GxE in this way is constrained to environments that fall into natural groupings (e.g., urban/rural residency) or forces researchers to create groups based on environments that may actually be continuous in nature (e.g., socioeconomic status). Moreover, stratification becomes impractical if the moderator is characterised by many levels and will effectively reduce the sample size, especially if the moderator is not shared between twins. Second, the use of heritability assumes equal variance across strata, whereas what is of interest in GxE analyses is changes in the magnitude of genetic and environmental effects, not only the proportion. Third, this approach cannot be used to disentangle gene-environment interaction and gene-environment correlation in a single analysis. For example, if individuals in a certain environment show greater genetic influence, this could be attributable to either (a) the environment moderating the effects of certain genes, or (b) certain trait-influencing genetic factors being more likely to be present in that environment.

Newer biometric moderation models have made it possible to examine moderation of the genetic and environmental influences on a trait along the full spectrum of the environmental measure (Purcell, 2002; van der Sluis et al., 2012). These newer biometric moderation models have several advantages over previous methods to study GxE in family datasets: (1) these models make it possible to formally test the presence/absence of interaction using validated fit indices, (2) it is possible to account for gene-environment correlation by modelling the main effect of the moderator on the phenotype, and (3) it is possible to examine moderation of the trait along the full spectrum of the moderator.

1.4.2.3. Theoretical GxE frameworks

Table 3 presents a summary of four main theoretical models that have been used to understand the nature of GxE: the diathesis stress, differential susceptibility, bioecological (social compensation), and social control/push models. In this section, I provide a brief description of each of the primary theories and relevant empirical examples from prior literature. Note that while these models are presented as distinct, it is entirely possible that more than one model may be needed to explain the pattern of GxE observed for a particular combination of moderator and phenotype. That is, these models should be regarded as useful

heuristics that can be used to interpret the effects identified in GxE studies, however, findings from any one specific study may suggest the plausibility of more than one framework.

Table 3. Summary of conceptual and theoretical models underlying gene-environment interaction (GxE) studies

Theory	Brief summary	Original article	Example of empirical article
Diathesis-stress	A predisposition for the trait (i.e., diathesis), in the form of premorbid risk factors (e.g., genetic susceptibility) lies dormant until triggered by an environmental stressor	Monroe & Simons (1991)	South & Krueger (2008) The study investigated the aetiology of the association between marital quality and internalising problems. Results showed that genetic effects on internalising problems increased as marital quality worsened, suggesting that those with a genetic predisposition to internalising syndromes may be more likely to express this predisposition in the context of a dissatisfying marriage.
Differential susceptibility	Individual differences in plasticity to the environment, with some people being more susceptible to (i.e., genetically influenced by) the effect of both positive and negative environments	Belsky & Pluess (2009)	South & Krueger (2013) The study investigated the nature of the association between martial satisfaction and physical health. Findings showed that genetic influences on physical health were greatest at both high and low levels of marital satisfaction, with lowest levels of heritability estimated for those at average level of marital satisfaction.
Bioecological model	Genetic influences are maximised in stable and adaptive environments that permit positive an stable interactions (proximal processes) between individuals and their environment, enabling them to reach their genetic potential	Bronfenbrenner & Ceci (1994)	Turkheimer et al. (2003) Seminal study that examined variation in the heritability of children's IQ as a function of family SES. Findings from this study showed that the heritability of IQ was higher among children in families with high SES compared to those with low SES.
Social control or social push	Genetic influences are filtered or buffered in certain environmental contexts; social control: social norms and structural constraints	Shanahan & Hofer (2005)	Boardman et al. (2010) The study investigated trends in the heritability of smoking across different birth cohorts. Results

	showed strong genetic influences on smoking for those born prior to legislation prohibiting smoking in public places, but negligible influences for those born subsequent to legislation efforts aimed at reducing smoking.

Source: Adapted from South et al. (2017) and Mills (2022)

The diathesis-stress model posits that genetic susceptibility (i.e., diathesis) for a trait lies dormant until it is triggered by an environmental stressor (Monroe & Simons, 1991). In this type of interaction, genetic influences become more apparent (i.e., explain a greater proportion of the phenotypic variance) in the presence of negative environmental conditions. In the context of twin studies, the diathesis-stress model would be supported if higher heritability estimates were observed among individuals in less favourable environments compared to those in more advantaged environments. This form of interaction (whereby genetic differences are enhanced in adverse environments) has been demonstrated as a function of several environmental stressors, such as poor marital quality, parental negativity and stressful life events (Dick, 2011; South & Krueger, 2008). To my knowledge, only one twin study investigating the effect of SES on the aetiology of depression symptoms found evidence to support the diathesis-stress model (Strachan et al., 2017). This study showed that genetic influences on depression symptoms were greater among individuals who lived in socioeconomically disadvantaged neighbourhoods compared to those who resided in more advantaged neighbourhoods.

Whereas the diathesis-stress model focuses primarily on negative environmental influences, the differential susceptibility model posits that individuals differ in plasticity, with some individuals more genetically susceptible to the effects of both favourable and adverse environments (Belsky & Pluess, 2009). In other words, individuals who are most sensitive to negative environmental influences may also be those who are most receptive to positive environmental influences.

The bioecological (or social compensation) model proposes that genetic influences are maximised in stable, adaptive and often more socioeconomically advantaged environments,

which enables individuals to reach their genetic potential. This form of interaction was supported in a seminal study that examined whether the heritability of children's IQ changes as a function of family SES (Turkheimer et al., 2003). Findings from this study showed that the heritability of IQ was higher in families with high SES compared to in families with low SES. An explanation for this finding can be linked to the Scarr-Rowe hypothesis that children in more advantaged families have higher resources that allows them to realise their genetic potential (Rowe et al., 1999; Scarr-Salapatek, 1971). In another example, studies have shown higher heritability of educational attainment following social and/or policy reforms that led to more equal opportunities (Heath et al., 1985; Rimfeld et al., 2018).

While the majority of studies reporting this form of interaction have been in the domain of cognitive ability and achievement, there is some evidence for this interaction as a possible mechanism underlying the association between socioeconomic status and mental health. Twin studies of emotional and behavioural problems in children have reported lower heritability and higher shared environmental influences (influences that make family members more similar to one another) in low SES families compared to high SES families (Burt et al., 2016, 2020; Hendriks et al., 2019; Middeldorp et al., 2014; Turkheimer et al., 2003). In turn, genetic influences were higher among children from high SES families compared to children from low SES families. Similarly, one study in adult twins found that non-shared environmental influences on internalising symptoms were greater at lower levels of income (South & Krueger, 2011). One interpretation of these findings is that at lower levels of SES, which are often characterised by multiple environmental risk factors (Braveman & Gottlieb, 2014; Wilkinson et al., 2003), genetic differences may be diminished as a result of greater variation in the social environment that confers risk for psychopathology. Higher heritability estimates among individuals with high SES corresponds to the notion that advantageous and/or low stress environments may amplify genetic differences underlying mental health problems.

The social control or social push model proposes that genetic influences are filtered or buffered by particular social environments, such as social norms (e.g., religious norms) or structural constraints (e.g., regional residency, alcohol taxation) that limit individual decision-making ability (Shanahan & Hofer, 2005). Thus, these factors serve as a control over the ability

to express genetic predispositions, and accordingly, reduce the degree to which genetic factors influence psychiatric and behavioural traits at the population level. For example, the heritability of smoking was found to be significantly reduced following the introduction of legislation prohibiting smoking in public places in the US (Boardman et al., 2010). Conversely, protective factors in the social environment may also buffer genetic predispositions towards mental health problems. For example, twin studies investigating the moderating effect of relationship status have shown that genetic liability for depression symptoms and alcohol consumption is decreased among individuals in committed relationships (Barr et al., 2017; Heath et al., 1998). This suggests that having a romantic partner may act as a protective 'social control' factor in reducing the impact of liability to psychiatric symptoms and substance misuse.

1.5. Challenges and biases in studying gene-environment interplay

1.5.1. Confounding by gene-environment correlation

A shortcoming of previous twin studies examining the moderating effect of family SES is that they do not model genetic overlap between SES and child psychopathology. This is an issue given the evidence for such genetic overlap from genome-wide genotype studies and extended family studies (Krapohl et al., 2017; Krapohl & Plomin, 2016; Torvik et al., 2020; Trzaskowski et al., 2014). The reason genetic overlap has not been modelled previously is that child twin data cannot be used to estimate genetic influences on family socioeconomic indicators because these variables do not differ between twin children within the same family. Thus, the genetic covariance between family SES and child outcomes cannot be estimated (Purcell, 2002; Rijsdijk & Sham, 2002). Failing to account for gene-environment correlation can lead to inflated signals of interaction and biased estimates (Purcell, 2002; van der Sluis et al., 2012).

Previous GxE twin studies examining the relationship between family SES and child psychopathology have controlled for the main effect of SES (moderator) on child outcomes prior to examining gene-environment interactions to account for inflation of test statistics due to rGE (Purcell, 2002; van der Sluis et al., 2012). As such, these prior studies specifically focus on whether family SES moderates the aetiological influences *unique* to child psychopathology.

However, this approach limits the investigation of the nature of the covariance between family SES and child psychopathology, as well as the moderation of shared variance. Given that family SES contributes to variance in child mental health (Krapohl et al., 2017; Torvik et al., 2020; Trzaskowski et al., 2014), exclusion of the common variance between them may limit our understanding of the underlying mechanisms through which socioeconomic disadvantage affects child mental health.

1.5.2. Limited geographic and demographic diversity in genetically informed research

A second issue undermining research is the lack of geographic and demographic diversity in much of the data currently used for genetically informed research. Within the literature discussed so far, analyses have been limited by a focus on specific family types (e.g., twins), often from a select subgroup of the population, and from high-income populations (Henrich et al., 2010). To date, twin data have been largely restricted to high-income populations, of which most studies were based on US twin samples (34%) (Polderman et al., 2015). Samples from South America, Africa and Asia were highly under-represented, constituting only 5.7% (combined) of studies. Limited representation for many populations also poses a significant challenge for genome-wide association studies, which have been largely conducted on European ancestry (Gurdasani et al., 2019).

Bias towards Western populations is problematic because estimates of genetic and environmental sources of individual differences are specific to a population at a particular time and have been shown to vary cross-culturally (J. Chen et al., 2014; Hur, 2008; Samuelsson et al., 2005; Selita & Kovas, 2019; Zavos et al., 2020). For example, higher heritability estimates of educational attainment have been observed in countries where the educational curriculum is highly standardised, because standardisation reduces environmental differences between schools (Samuelsson et al., 2005). Twin studies have also found some evidence to suggest that the aetiology of mental health problems is different in LMIC than in high-income countries (Hur, 2008; Zavos et al., 2020). For example, sex differences in the aetiology of depression have been observed in non-Western populations, with men showing low heritability and women showing moderate heritability in both Sri Lanka and South Korea (Hur, 2008; Zavos et

al., 2020). One explanation for the differences in the origins of depression across culture and gender is that the social environment is variable both within and across populations.

Limited representation of research in LMIC populations poses a significant challenge for understanding the mechanisms underpinning socioeconomic inequalities in mental health outcomes in these populations. Patterns of gene-environment interplay observed in high-income populations may not reflect those at play in other populations with different social stratification processes (e.g., the socioeconomic distribution may be greater and different socioeconomic factors may be more or less important in LMIC versus HIC). Thus, the results from GxE research in HIC may not be generalisable to LMIC populations. This could result in erroneous conclusions about the mechanisms through which socioeconomic factors impact mental health outcomes in these populations, which could have important implications for the development of interventions and treatments for individuals who are at risk for mental health disorders.

1.5.3. Non-representative samples and selection bias

Another related issue is that much of the research on socioeconomic disparities in mental health is based on non-representative samples (Munafò et al., 2018; Nohr et al., 2006; Nohr & Liew, 2018). Sociodemographic, lifestyle and health factors have been associated with the likelihood of individuals becoming or remaining participants in a study. Those who are less advantaged and less healthy are often under-represented in studies, leading to samples that are generally 'healthier and wealthier' than the intended study population (Dupuis et al., 2019; Goldberg et al., 2001; Graaf et al., 2000; Hara et al., 2002; Howe et al., 2012; Lamers et al., 2012; Nohr & Liew, 2018). This is a concern because selective participation, both at recruitment and during subsequent follow-ups, can introduce selection bias and compromise both the internal and external validity of study results (Munafò et al., 2018). Selection biases can be influenced by factors operating at a global (e.g., inequality in research funding and infrastructure may lead to the clustering of cohorts in developed countries), national (e.g., cultural and societal factors may lead to gender and racial biases in study participation), institutional (e.g., lack of diversity among scientists who design and implement research

protocol may lead to biases in research) and individual (e.g., greater exposure to life stressors may lead to lower capacity to engage in research) level.

Selection bias is common in scientific research that relies on voluntary participation. It occurs when there are systematic differences between individuals who participate and those who do not. Selection can occur at various stages in a study, including at recruitment, at follow-up, or if a non-random subsample is selected for further investigation (e.g., based on exposure to a specific risk factor). It is widely acknowledged that selection bias substantially influences prevalence estimates, often towards underestimating prevalence rates of mental health problems. In contrast, it has often been assumed that selection bias only has a minimal impact on observed associations (Munafò et al., 2018). Recent methodological reviews, as well as empirical studies using data from various population-based studies (e.g., ALSPAC, MoBa, UK Biobank), suggests that this assumption is problematic in many circumstances (Adams et al., 2020; Biele et al., 2019; Munafò et al., 2018; Taylor et al., 2018). When there are differences between participants and nonparticipants that relate to the risk factors and outcomes being investigated, associations between risk and outcome may become biased (Munafò et al., 2018). In other words, selection can bias associations between variables that influence participation in a study. In such circumstances, this can induce associations where there is no causal effect, attenuate true causal effects, or reverse the sign of a causal effect. For example, research using data from the Avon Longitudinal Study of Parents and Children (ALSPAC) showed that attrition from cohort studies can underestimate socioeconomic inequalities in various health outcomes, and the degree of bias may worsen with increasing drop-out rates (Howe et al., 2013).

Recent genetic association studies have explored factors that influence selective participation in order to better understand biases that may affect study samples. Genome-wide association studies have identified a number of loci associated with continued participation in cohorts such as ALSPAC (Taylor et al., 2018), UK Biobank (Adams et al., 2020), and the Norwegian Mother, Father and Child Cohort Study (MoBa) (Ask et al., *in prep*; Biele et al., 2019), with estimates of SNP-based heritability ranging from 6.5% to 32%. Genetic overlap has been observed between continued study participation and several traits. For example, continued participation in ALSPAC was linked to polygenic scores for higher education, agreeableness,

and openness, while attrition was associated with higher polygenic scores for BMI, smoking, neuroticism, schizophrenia, ADHD, and depression (Taylor et al., 2018). In UK Biobank, continued participation in a mental health questionnaire was found to be genetically correlated with better health, higher education, and lower psychiatric disorder rates (Adams et al., 2020). In a separate study, strong positive genetic correlations were found between continued participation in ALSPAC and participation in three of the optional UK Biobank follow-up surveys (Tyrrell et al., 2020), indicating that similar genetic factors are associated with participation at follow-up assessments across the two studies.

Studies have also shown that selection bias can affect the magnitude of associations between polygenic scores and phenotypic outcomes, with conditioning on selection potentially inducing associations that are not present in the population. Munafò et al. (2018) found that polygenic risk for smoking was associated with maternal education in a selected sub-study of ALSPAC, but not in the full cohort. Similarly, Taylor et al. (2018) showed that associations between polygenic score for education and being an ever smoker, and between the education polygenic score and BMI, were weaker in a subsample of those who attended the most recent clinic compared to the full genetic sample. These studies demonstrate the potential for bias in genetic analyses when studying selected subsamples based on the availability of follow-up data.

Selection bias poses several challenges for research seeking to understand the causal processes underpinning socioeconomic inequalities in mental health. First, studies that over-represent individuals with higher socioeconomic status and better mental health may attenuate the strength of the association between these factors and outcomes. Individuals with higher SES and better mental health tend to have greater access to resources and support that can buffer the impact of social stressors (e.g., financial insecurity) on mental health (Braveman & Gottlieb, 2014; Wilkinson et al., 2003). Thus, disproportionate representation of study participants who are 'healthier and wealthier' than the larger population can lead to an underestimation of the impact of low socioeconomic status on mental health outcomes (e.g., Howe et al., 2013) and may result in misleading implications for the design and implementation of interventions and policy.

The effects of genetic and environmental factors on mental health outcomes may be biased by selection. If the study sample is not representative of the broader population, the genetic (and environmental) variation in the sample may not accurately reflect the genetic variation in the broader population. Consequently, selective participation may bias heritability estimates, as well as estimates of aetiological overlap between socioeconomic factors and psychiatric traits (Akimova et al., 2021; Marees et al., 2021).

Similarly, selection bias has the potential to influence GxE analyses in several ways. First, selection bias can limit the generalisability of GxE study findings. If the study sample is not representative of the general population, it may not be possible to generalise the results to the broader population. Second, selection bias can reduce variability in the exposure and outcome in the study sample. This reduces the power to detect GxE interactions and/or lead to an underestimation of the true interaction effect (Dick, 2011; Murray et al., 2016; Purcell, 2002; D. Reiss et al., 2013). For example, if the study sample is over-representative of certain groups, such as those from higher socioeconomic backgrounds, estimates of GxE interactions may be biased towards protective effects of the social environment, as individuals from higher socioeconomic backgrounds often have lower exposure to social environmental risk factors (leading to reduced variability in environmental exposures) and/or more resources to mitigate the effects of environmental risk factors. Third, selection bias can lead to erroneous conclusions about the mechanisms by which genetic and environmental influences moderate each other, as variation in the distributions of individuals along both the exposure and outcome variables can alter the shape and significance of interactions (Dick, 2011; Murray et al., 2016; Purcell, 2002; D. Reiss et al., 2013). Failing to capture the full spectrum of socioeconomic background and/or mental health can preclude researchers from detecting interaction effects at the lowest or highest levels of exposure and/or psychopathology. As a result, this could have important implications for designing and implementing new forms of prevention and intervention.

1.5.4. Summarising gaps in the literature

The research published to date suggests that both social causation (i.e., gene-environment interaction) and social selection (i.e., gene-environment correlation) processes contribute to

observed associations between socioeconomic conditions and mental health outcomes (Dohrenwend et al., 1992; Hoffmann et al., 2018; Jin et al., 2020; Lund & Cois, 2018). However, significant gaps remain in our understanding. Few studies have explored the moderating effects of socioeconomic status indicators on the aetiology of mental health outcomes using a behavioural genetic approach. Even fewer have explored potential differences in the effects of individual socioeconomic indicators. Further, the use of child twin data in prior GxE studies has not allowed researchers to test for gene-environment interaction in the presence of gene-environment correlation (Purcell, 2002; van der Sluis et al., 2012). Crucially, no studies have investigated gene-environment interplay processes underpinning the relationship between socioeconomic status and mental health outcomes in lower- and middle-income populations. Finally, questions remain as to our ability to detect the true effects of socioeconomic conditions on mental health, given the potential for bias in phenotypic and genetic analyses when studying samples characterised by self-selection.

1.6. Introducing aims for this thesis

1.6.1. Improving understanding of the mechanisms underlying socioeconomic disparities in children's mental health: gene-environment interaction in the presence of gene-environment correlation

As outlined in previous sections, children from socioeconomically disadvantaged backgrounds have an increased risk of experiencing mental health difficulties (F. Reiss, 2013). However, gaps remain in our understanding as to the mechanisms underlying these observed associations. There is some evidence from twin studies that family socioeconomic conditions influence the relative importance of genetic and environmental influences on child emotional and behavioural problems (Burt et al., 2016, 2020; Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). However, it is also known that various socioeconomic indicators (e.g., income level, educational attainment) are themselves under genetic influence (Hill et al., 2019; Krapohl et al., 2017; Trzaskowski et al., 2014) and previous gene-environment interaction studies have not explored the potential genetic overlap between family SES and child mental health. Exclusion of the common variance between family SES and child outcomes limits our understanding as to the nature of their covariance, as well as the moderation of the shared variance.

In this thesis, I aim to contribute new knowledge on the role that family socioeconomic factors play in shaping the origins of child emotional and behavioural problems. In chapter 2, I apply a novel approach using extended family data to investigate the moderation of aetiologic influences on child emotional and behavioural problems by parental socioeconomic status in the presence of gene-environment correlation. To my knowledge, this is first study to test for gene-environment interaction in the presence of gene-environment correlation for environmental moderators that are necessarily shared between children growing up in the same family. Further, inconsistent findings from the literature may in part be explained by the use of different socioeconomic indicators across studies (Duncan & Magnuson, 2003; Geyer et al., 2006; F. Reiss, 2013). This suggests that the presence and pattern of moderation effects may vary depending on the measure used to index family SES. I build on this work to examine the moderating effects of individual socioeconomic indicators, namely income and education, separately for both mothers and fathers.

1.6.2. Improving understanding of the mechanisms underpinning socioeconomic disparities in mental health in low- and middle-income populations

Within the literature discussed so far, GxE analyses have been limited to samples drawn from high-income countries. Thus, findings from these studies may not generalise to countries with greater economic disparity, higher levels of poverty, and other differences in social stratification processes (e.g., gender norms). There is limited representation of genetically informative research conducted in lower- and middle-income populations (Polderman et al., 2015), even as the greatest health disparities and disease burden are observed in LMIC (Maselko, 2017; Saxena et al., 2006). This limits our understanding of the processes underpinning the relationship between SES and mental health in these populations.

In chapter 3, I make use of the rich data available from the Colombo Twins and Singletons Study (CoTASS) to investigate individual differences in socioeconomic factors and their influence on depression in Sri Lanka. Using structural equation modelling, I explore genetic and environmental influences on individual socioeconomic indicators and depression symptoms and moderation of aetiological influences on depression as a function of these

socioeconomic indicators. In this study, I aim to improve our understanding of the mechanisms underlying the relationship between socioeconomic status and depression symptoms in a South Asian population.

1.6.3. Considering selection bias in family data

As outlined in the previous section, many cohort and population-based studies are based on non-representative samples, with under-representation of less advantaged and less healthy individuals. This often leads to samples that are generally 'healthier and wealthier' than the intended study population, raising concerns about the potential for selection bias and compromising the validity of study results (Munafò et al., 2018). Attrition from cohort studies has the potential to introduce bias in prevalence estimates and impact associations between variables, potentially distorting causal relationships (Howe et al., 2013; Munafò et al., 2018). Recent genetic association studies have explored factors influencing selective participation and have identified genetic overlap between continued study participation and various traits (Adams et al., 2020; Biele et al., 2019; Taylor et al., 2018; Tyrrell et al., 2020). Studies have also shown that selection bias can affect associations between polygenic scores and phenotypic outcomes. Selection bias poses challenges for research on socioeconomic disparities in mental health. Over-representing individuals with higher socioeconomic status and better mental health may attenuate the association between these factors and outcomes, potentially leading to an underestimation of the impact of low socioeconomic status on mental health. Selective participation may also bias heritability estimates and influence GxE analyses by limiting generalisability, reducing variability in exposure and outcome, and altering the shape and significance of interactions. This highlights the need for addressing selection bias in research studies and its potential implications for understanding the mechanisms underlying the observed SES – mental health relationship.

In chapter 4, I aim to contribute new insights into the impact of selection bias on phenotypic and genetic correlations within the Norwegian Mother, Father and Child Birth Cohort Study (MoBa). I seek to expand on previous research by examining the factors associated with continued participation in MoBa and their links to various baseline variables, socioeconomic factors, and maternal and offspring outcomes. Specifically, I focus on the potential bias arising

from selection on the associations between maternal educational attainment or income rank and outcomes related to maternal and offspring internalising and externalising symptoms.

To achieve this, I employ biometric model fitting to estimate genetic and environmental influences associated with maternal continued participation and investigate genetic and environmental overlap with maternal educational attainment and internalising symptoms. I also explore potential biases resulting from selection on the genetic and environmental correlations between maternal educational attainment and internalising symptoms. By undertaking these analyses, I aim to contribute to the existing literature on selection bias in population-based cohort studies, with a particular focus on the relationship between maternal socioeconomic factors and internalising problems in MoBa. With these findings, I aim to provide insights into the potential biases introduced by selective participation and enhance understanding of the associations between maternal participation, educational attainment, internalising symptoms, and offspring outcomes.

1.7. Summarising the aims and structure of thesis

The aim of this thesis is to advance our understanding of the mechanisms underlying socioeconomic disparities in children's and adults' mental health. The thesis is structured into three main chapters, focusing on gene-environment interactions and the impact of selection bias in family data:

Chapter 2 investigates the role of family socioeconomic factors in shaping the origins of child emotional and behavioural problems. Using extended family data from MoBa, I apply a novel approach to examine how parental socioeconomic factors moderate aetiological influences on these mental health outcomes in the presence of gene-environment correlation. By addressing gaps in the literature and considering different measures of family socioeconomic status, this study aims to enhance our understanding of the complex relationships between family socioeconomic conditions and children’s mental health.

Chapter 3 focuses on understanding the mechanisms underlying socioeconomic disparities in mental health in low- and middle-income populations. Using data from the Colombo Twins and Singletons Study (CoTASS) in Sri Lanka, I examine the influence of individual

socioeconomic factors on depression symptoms in adults. Using a twin and singleton design, I investigate the genetic and environmental influences on individual socioeconomic indicators and their moderation of the aetiological influences on depression. This study aims to contribute to the limited representation of genetically informative research conducted in these populations and provide insights into the relationship between socioeconomic status and mental health outcomes in a South Asian context.

Chapter 4 examines the impact of selection bias on phenotypic and genetic correlations within population-based cohort studies, focusing on the MoBa study. By examining the factors associated with continued participation in MoBa and their links to various baseline variables, socioeconomic factors, and mental health outcomes, I aim to assess the potential bias arising from selection in this cohort. Specifically, I investigate the associations between maternal educational attainment or income rank and internalising and externalising symptoms in mothers and offspring. Employing biometric model fitting, I estimate the heritability of continued participation and explore the genetic and environmental overlap with maternal socioeconomic indicators and internalising outcomes. Through these analyses, I aim to contribute to the literature on selection bias in population-based cohort studies and enhance our understanding of the associations between socioeconomic factors and mental health disparities.

Overall, this thesis aims to advance our understanding of socioeconomic disparities in mental health outcomes, including both children and adults, by examining gene-environment interactions and addressing the impact of selection bias. The findings from this research have implications for interventions and policies aimed at reducing mental health disparities and promoting well-being across diverse populations.

1.8. References

Achenbach, T., & Edelbrock, C. (1983, June 1). *Manual for the Child: Behavior Checklist and Revised Child Behavior Profile*. https://www.semanticscholar.org/paper/Manual-for-the-Child%3A-Behavior-Checklist-and-Child-Achenbach-Edelbrock/408859045620a1a00dc2ad253a9c799cdef51eff

Adams, M. J., Hill, W. D., Howard, D. M., Dashti, H. S., Davis, K. A. S., Campbell, A., Clarke, T.-K., Deary, I. J., Hayward, C., Porteous, D., Hotopf, M., & McIntosh, A. M. (2020). Factors associated with sharing e-mail information and mental health survey participation in large population cohorts. *International Journal of Epidemiology*, *49*(2), 410–421. https://doi.org/10.1093/ije/dyz134

Akee, R., Copeland, W., Costello, E. J., & Simeonova, E. (2018). How Does Household Income Affect Child Personality Traits and Behaviors? *American Economic Review*, *108*(3), 775–827. https://doi.org/10.1257/aer.20160133

Akimova, E. T., Breen, R., Brazel, D. M., & Mills, M. C. (2021). Gene-environment dependencies lead to collider bias in models with polygenic scores. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-89020-x

Alegría, M., NeMoyer, A., Falgas, I., Wang, Y., & Alvarez, K. (2018). Social Determinants of Mental Health: Where We Are and Where We Need to Go. *Current Psychiatry Reports*, *20*(11), 95. https://doi.org/10.1007/s11920-018-0969-9

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Andreassen, O. A., Hindley, G. F. L., Frei, O., & Smeland, O. B. (2023). New insights from the last decade of research in psychiatric genetics: Discoveries, challenges and clinical implications. *World Psychiatry*, *22*(1), 4–24. https://doi.org/10.1002/wps.21034

Angel, S., Disslbacher, F., Humer, S., & Schnetzer, M. (2019). What did you really earn last year?: Explaining measurement error in survey income data. *Journal of the Royal Statistical Society Series A*, *182*(4), 1411–1437.

Antonoplis, S. (2023). Studying Socioeconomic Status: Conceptual Problems and an Alternative Path Forward. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *18*(2), 275–292. https://doi.org/10.1177/17456916221093615

Araya, R., Lewis, G., Rojas, G., & Fritsch, R. (2003). Education and income: Which is more important for mental health? *Journal of Epidemiology & Community Health*, *57*(7), 501–505. https://doi.org/10.1136/jech.57.7.501

Arias, D., Saxena, S., & Verguet, S. (2022). Quantifying the global burden of mental disorders and their economic value. *EClinicalMedicine*, *54*. https://doi.org/10.1016/j.eclinm.2022.101675

Ask, H., Corfield, E., Biele, G., Tilling, K., McAdams, T. A., Badini, I., Davies, N. M., Davey Smith, G., Richardson, T. G., Cornish, R., Tesli, M., Askeland, R. B., Magnus, P., Andreassen, O. A., Njølstad, P. R., Ystrøm, E., Stoltenberg, C., Reichborn-Kjennerud, T., Hannigan, L. J., & Havdahl, A. (in prep). *The genomic basis of continued participation in the Norwegian Mother, Father and Child Cohort Study*.

Assary, E., Vincent, J. P., Keers, R., & Pluess, M. (2018). Gene-environment interaction and psychiatric disorders: Review and future directions. *Seminars in Cell & Developmental Biology*, *77*, 133–143. https://doi.org/10.1016/j.semcdb.2017.10.016

Ball, H. A., Siribaddana, S. H., Sumathipala, A., Kovas, Y., Glozier, N., McGuffin, P., & Hotopf, M. (2010). Environmental exposures and their genetic or environmental contribution to depression and fatigue: A twin study in Sri Lanka. *BMC Psychiatry*, *10*, 13. https://doi.org/10.1186/1471-244X-10-13

Barr, P. B., Salvatore, J. E., Maes, H. H., Korhonen, T., Latvala, A., Aliev, F., Viken, R., Rose, R. J., Kaprio, J., & Dick, D. M. (2017). Social Relationships Moderate Genetic Influences on Heavy Drinking in Young Adulthood. *Journal of Studies on Alcohol and Drugs*, *78*(6), 817–826. https://doi.org/10.15288/jsad.2017.78.817

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, *135*(6), 885–908. https://doi.org/10.1037/a0017376

Biele, G., Gustavson, K., Czajkowski, N. O., Nilsen, R. M., Reichborn-Kjennerud, T., Magnus, P. M., Stoltenberg, C., & Aase, H. (2019). Bias from self selection and loss to follow-up in prospective cohort studies. *European Journal of Epidemiology*, *34*(10), 927–938. https://doi.org/10.1007/s10654-019-00550-1

Boardman, J. D., Blalock, C. L., & Pampel, F. C. (2010). Trends in the Genetic Influences on Smoking. *Journal of Health and Social Behavior*, *51*(1), 108–123. https://doi.org/10.1177/0022146509361195

Braveman, P., & Gottlieb, L. (2014). The Social Determinants of Health: It's Time to Consider the Causes of the Causes. *Public Health Reports*, *129*(Suppl 2), 19–31.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, *101*(4), 568–586. https://doi.org/10.1037/0033-295x.101.4.568

Burns, J. K. (2015). Poverty, inequality and a political economy of mental health. *Epidemiology and Psychiatric Sciences*, *24*(2), 107–113. https://doi.org/10.1017/S2045796015000086

Burt, S. A., Klump, K. L., Gorman-Smith, D., & Neiderhiser, J. M. (2016). Neighborhood Disadvantage Alters the Origins of Children's Nonaggressive Conduct Problems. *Clinical Psychological Science : A Journal of the Association for Psychological Science*, *4*(3), 511–526. https://doi.org/10.1177/2167702615618164

Burt, S. A., Pearson, A. L., Carroll, S., Klump, K. L., & Neiderhiser, J. M. (2020). Child Antisocial Behavior Is more Environmental in Origin in Disadvantaged Neighborhoods: Evidence Across Residents' Perceptions and Geographic Scales in Two Samples. *Journal of Abnormal Child Psychology*, *48*(2), 265–276. https://doi.org/10.1007/s10802-019-00587-6

Chen, J., Li, X., Natsuaki, M. N., Leve, L. D., & Harold, G. T. (2014). Genetic and Environmental Influences on Depressive Symptoms in Chinese Adolescents. *Behavior Genetics*, *44*(1), 36–44. https://doi.org/10.1007/s10519-013-9632-8

Chen, Q., Brikell, I., Lichtenstein, P., Serlachius, E., Kuja-Halkola, R., Sandin, S., & Larsson, H. (2017). Familial aggregation of attention-deficit/hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, *58*(3), 231–239. https://doi.org/10.1111/jcpp.12616

Clarke, C., Bonnet, J., Flores, M., & Thévenon, O. (2022). *The economic costs of childhood socio-economic disadvantage in European OECD countries*. OECD. https://doi.org/10.1787/8c0c66b9-en

Conway, D. I., McMahon, A. D., Brown, D., & Leyland, A. H. (2019). Measuring socioeconomic status and inequalities. In S. Vaccarella, J. Lortet-Tieulent, R. Saracci, D. I. Conway, K. Straif, & C. P. Wild (Eds.), *Reducing social inequalities in cancer: Evidence and priorities for research*. International Agency for Research on Cancer. http://www.ncbi.nlm.nih.gov/books/NBK566205/

Cooper, S., Lund, C., & Kakuma, R. (2012). The measurement of poverty in psychiatric epidemiology in LMICs: Critical review and recommendations. *Social Psychiatry and*

Psychiatric Epidemiology, *47*(9), 1499–1516. https://doi.org/10.1007/s00127-011-0457-6

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships Between Poverty and PsychopathologyA Natural Experiment. *JAMA*, *290*(15), 2023–2029. https://doi.org/10.1001/jama.290.15.2023

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature*, *47*(1), 87–122. https://doi.org/10.1257/jel.47.1.87

Daré, L. O., Bruand, P.-E., Gérard, D., Marin, B., Lameyre, V., Boumédiène, F., & Preux, P.-M. (2019). Co-morbidities of mental disorders and chronic physical diseases in developing and emerging countries: A meta-analysis. *BMC Public Health*, *19*(1), 304. https://doi.org/10.1186/s12889-019-6623-6

Darin-Mattsson, A., Fors, S., & Kåreholt, I. (2017). Different indicators of socioeconomic status and their relative importance as determinants of health in old age. *International Journal for Equity in Health*, *16*(1), 173. https://doi.org/10.1186/s12939-017-0670-3

Dick, D. M. (2011). Gene-Environment Interaction in Psychological Traits and Disorders. *Annual Review of Clinical Psychology*, *7*, 383–409. https://doi.org/10.1146/annurev-clinpsy-032210-104518

Diemer, M. A., Mistry, R. S., Wadsworth, M. E., López, I., & Reimers, F. (2013). Best Practices in Conceptualizing and Measuring Social Class in Psychological Research. *Analyses of Social Issues and Public Policy*, *13*(1), 77–113. https://doi.org/10.1111/asap.12001

Dohrenwend, B. P., Levav, I., Shrout, P. E., Schwartz, S., Naveh, G., Link, B. G., Skodol, A. E., & Stueve, A. (1992). Socioeconomic status and psychiatric disorders: The causation-selection issue. *Science (New York, N.Y.)*, *255*(5047), 946–952. https://doi.org/10.1126/science.1546291

Domènech-Abella, J., Mundó, J., Leonardi, M., Chatterji, S., Tobiasz-Adamczyk, B., Koskinen, S., Ayuso-Mateos, J. L., & Haro, J. M. (2018). The association between socioeconomic status and depression among older adults in Finland, Poland and Spain: A comparative cross-sectional study of distinct measures and pathways. *Journal of Affective Disorders*, *241*, 311–318. https://doi.org/10.1016/j.jad.2018.08.077

Duncan, G. J., & Magnuson, K. A. (2003). Off with Hollingshead: Socioeconomic resources, parenting, and child development. In *Socioeconomic status, parenting, and child development* (pp. 83–106). Lawrence Erlbaum Associates Publishers.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How Much Does Childhood Poverty Affect the Life Chances of Children? *American Sociological Review*, *63*(3), 406–423. https://doi.org/10.2307/2657556

Dupuis, M., Strippoli, M.-P. F., Gholam-Rezaee, M., Preisig, M., & Vandeleur, C. L. (2019). Mental disorders, attrition at follow-up, and questionnaire non-completion in epidemiologic research. Illustrations from the CoLaus|PsyCoLaus study. *International Journal of Methods in Psychiatric Research*, *28*(4), e1805. https://doi.org/10.1002/mpr.1805

Eriksson, K., Lindvall, J., Helenius, O., & Ryve, A. (2021). Socioeconomic Status as a Multidimensional Predictor of Student Achievement in 77 Societies. *Frontiers in Education*, *6*. https://www.frontiersin.org/articles/10.3389/feduc.2021.731634

Festin, K., Thomas, K., Ekberg, J., & Kristenson, M. (2017). Choice of measure matters: A study of the relationship between socioeconomic status and psychosocial resources in a middle-aged normal population. *PLOS ONE*, *12*(8), e0178929. https://doi.org/10.1371/journal.pone.0178929

Fitzsimons, E., Goodman, A., Kelly, E., & Smith, J. P. (2017). Poverty dynamics and parental mental health: Determinants of childhood mental health in the UK. *Social Science & Medicine*, *175*, 43–51. https://doi.org/10.1016/j.socscimed.2016.12.040

Flaskerud, J. H., & DeLilly, C. R. (2012). Social Determinants of Health Status. *Issues in Mental Health Nursing*, *33*(7), 494–497. https://doi.org/10.3109/01612840.2012.662581

GBD 2019 Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet. Psychiatry*, *9*(2), 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3

Geyer, S., Hemström, Ö., Peter, R., & Vågerö, D. (2006). Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *Journal of Epidemiology & Community Health*, *60*(9), 804–810. https://doi.org/10.1136/jech.2005.041319

Goldberg, M., Chastang, J. F., Leclerc, A., Zins, M., Bonenfant, S., Bugel, I., Kaniewski, N., Schmaus, A., Niedhammer, I., Piciotti, M., Chevalier, A., Godard, C., & Imbernon, E. (2001). Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: A prospective study of the French GAZEL cohort and its target population. *American Journal of Epidemiology*, *154*(4), 373–384. https://doi.org/10.1093/aje/154.4.373

Graaf, R. de, Bijl, R. V., Smit, F., Ravelli, A., & Vollebergh, W. A. M. (2000). Psychiatric and Sociodemographic Predictors of Attrition in a Longitudinal Study The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *American Journal of Epidemiology*, *152*(11), 1039–1047. https://doi.org/10.1093/aje/152.11.1039

Gurdasani, D., Barroso, I., Zeggini, E., & Sandhu, M. S. (2019). Genomics of disease risk in globally diverse populations. *Nature Reviews Genetics*, *20*(9), Article 9. https://doi.org/10.1038/s41576-019-0144-0

Hakulinen, C., Mok, P. L. H., Horsdal, H. T., Pedersen, C. B., Mortensen, P. B., Agerbo, E., & Webb, R. T. (2020). Parental income as a marker for socioeconomic position during childhood and later risk of developing a secondary care-diagnosed mental disorder examined across the full diagnostic spectrum: A national cohort study. *BMC Medicine*, *18*(1), 323. https://doi.org/10.1186/s12916-020-01794-5

Hara, M., Sasaki, S., Sobue, T., Yamamoto, S., & Tsugane, S. (2002). Comparison of cause-specific mortality between respondents and nonrespondents in a population-based prospective study: Ten-year follow-up of JPHC Study Cohort I. *Journal of Clinical Epidemiology*, *55*(2), 150–156. https://doi.org/10.1016/S0895-4356(01)00431-0

Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Corey, L. A., Sundet, J., Magnus, P., & Nance, W. E. (1985). Education policy and the heritability of educational attainment. *Nature*, *314*(6013), Article 6013. https://doi.org/10.1038/314734a0

Heath, A. C., Eaves, L. J., & Martin, N. G. (1998). Interaction of marital status and genetic risk for symptoms of depression. *Twin Research and Human Genetics*, *1*(3), 119–122. https://doi.org/10.1375/twin.1.3.119

Hendriks, A. M., Finkenauer, C., Nivard, M. G., Van Beijsterveldt, C. E. M., Plomin, R. J., Boomsma, D. I., & Bartels, M. (2019). Comparing the genetic architecture of childhood behavioral problems across socioeconomic strata in the Netherlands and

the United Kingdom. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-019-01357-x

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61–83; discussion 83-135. https://doi.org/10.1017/S0140525X0999152X

Hettema, J. M., Neale, M. C., & Kendler, K. S. (2001). A Review and Meta-Analysis of the Genetic Epidemiology of Anxiety Disorders. *American Journal of Psychiatry*, *158*(10), 1568–1578. https://doi.org/10.1176/appi.ajp.158.10.1568

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

Hoebel, J., Maske, U. E., Zeeb, H., & Lampert, T. (2017). Social Inequalities and Depressive Symptoms in Adults: The Role of Objective and Subjective Socioeconomic Status. *PLOS ONE*, *12*(1), e0169764. https://doi.org/10.1371/journal.pone.0169764

Hoffmann, R., Kröger, H., & Pakpahan, E. (2018). Pathways between socioeconomic status and health: Does health selection or social causation dominate in Europe? *Advances in Life Course Research*, *36*, 23–36. https://doi.org/10.1016/j.alcr.2018.02.002

Houweling, T. A. J., Oude Groeniger, J., Jansen, P. W., van Lier, P., Horoz, N., Buil, M., & van Lenthe, F. J. (2022). Trajectories of socioeconomic inequality in early child development: A cohort analysis. *International Journal for Equity in Health*, *21*(1), 79. https://doi.org/10.1186/s12939-022-01675-8

Howe, L. D., Galobardes, B., Matijasevich, A., Gordon, D., Johnston, D., Onwujekwe, O., Patel, R., Webb, E. A., Lawlor, D. A., & Hargreaves, J. R. (2012). Measuring socio-economic position for epidemiological studies in low- and middle-income countries: A methods of measurement in epidemiology paper. *International Journal of Epidemiology*, *41*(3), 871–886. https://doi.org/10.1093/ije/dys037

Howe, L. D., Tilling, K., Galobardes, B., & Lawlor, D. A. (2013). Loss to follow-up in cohort studies: Bias in estimates of socioeconomic inequalities. *Epidemiology (Cambridge, Mass.)*, *24*(1), 1–9. https://doi.org/10.1097/EDE.0b013e31827623b1

Hur, Y.-M. (2008). Sex Differences in Genetic and Environmental Contributions to Depression Symptoms in South Korean Adolescent and Young Adult Twins. *Twin Research and Human Genetics*, *11*(3), 306–313. https://doi.org/10.1375/twin.11.3.306

Hyytinen, A., Ilmakunnas, P., Johansson, E., & Toivanen, O. (2019). Heritability of lifetime earnings. *The Journal of Economic Inequality*, *17*(3), 319–335. https://doi.org/10.1007/s10888-019-09413-x

Jaffee, S. R., Moffitt, T. E., Caspi, A., Fombonne, E., Poulton, R., & Martin, J. (2002). Differences in early childhood risk factors for juvenile-onset and adult-onset depression. *Archives of General Psychiatry*, *59*(3), 215–222. https://doi.org/10.1001/archpsyc.59.3.215

Jin, Y., Zhu, D., & He, P. (2020). Social causation or social selection? The longitudinal interrelationship between poverty and depressive symptoms in China. *Social Science & Medicine*, *249*, 112848. https://doi.org/10.1016/j.socscimed.2020.112848

Kalin, N. H. (2020). The Critical Relationship Between Anxiety and Depression. *American Journal of Psychiatry*, *177*(5), 365–367. https://doi.org/10.1176/appi.ajp.2020.20030305

Kendall, K. M., Assche, E. V., Andlauer, T. F. M., Choi, K. W., Luykx, J. J., Schulte, E. C., & Lu, Y. (2021). The genetic basis of major depression. *Psychological Medicine*, *51*(13), 2217–2230. https://doi.org/10.1017/S0033291721000441

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., Ustün, T. B., & Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) surveys. *Epidemiologia E Psichiatria Sociale*, *18*(1), 23–33. https://doi.org/10.1017/s1121189x00001421

Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Üstün, T. B. (2007). Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry*, *20*(4), 359. https://doi.org/10.1097/YCO.0b013e32816ebc8c

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H.-U. (1998). The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF). *International Journal of Methods in Psychiatric Research*, *7*(4), 171–185. https://doi.org/10.1002/mpr.47

Kinge, J. M., Øverland, S., Flatø, M., Dieleman, J., Røgeberg, O., Magnus, M. C., Evensen, M., Tesli, M., Skrondal, A., Stoltenberg, C., Vollset, S. E., Håberg, S., & Torvik, F. A. (2021). Parental income and mental disorders in children and adolescents: Prospective register-based study. *International Journal of Epidemiology*, *dyab066*. https://doi.org/10.1093/ije/dyab066

Kivimäki, M., Batty, G. D., Pentti, J., Shipley, M. J., Sipilä, P. N., Nyberg, S. T., Suominen, S. B., Oksanen, T., Stenholm, S., Virtanen, M., Marmot, M. G., Singh-Manoux, A., Brunner, E. J., Lindbohm, J. V., Ferrie, J. E., & Vahtera, J. (2020). Association between socioeconomic status and the development of mental and physical health conditions in adulthood: A multi-cohort study. *The Lancet Public Health*, *5*(3), e140–e149. https://doi.org/10.1016/S2468-2667(19)30248-8

Knopik, V. S., Neiderhiser, J. M., DeFries, J. C., & Plomin, R. (2016). *Behavioral Genetics* (7th ed.). Macmillan Learning.

Korous, K. M., Bradley, R. H., Luthar, S. S., Li, L., Levy, R., Cahill, K. M., & Rogers, C. R. (2022). Socioeconomic status and depressive symptoms: An individual-participant data meta-analysis on range restriction and measurement in the United States. *Journal of Affective Disorders*, *314*, 50–58. https://doi.org/10.1016/j.jad.2022.06.090

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

Krapohl, E., & Plomin, R. (2016). Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular Psychiatry*, *21*(3), 437–443. https://doi.org/10.1038/mp.2015.2

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine*, *146*(5), 317–325. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

Lamers, F., Hoogendoorn, A. W., Smit, J. H., van Dyck, R., Zitman, F. G., Nolen, W. A., & Penninx, B. W. (2012). Sociodemographic and psychiatric determinants of attrition in

the Netherlands Study of Depression and Anxiety (NESDA). *Comprehensive Psychiatry*, *53*(1), 63–70. https://doi.org/10.1016/j.comppsych.2011.01.011

Lansford, J. E., Malone, P. S., Tapanya, S., Tirado, L. M. U., Zelli, A., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., & Steinberg, L. (2019). Household Income Predicts Trajectories of Child Internalizing and Externalizing Behavior in High-, Middle-, and Low-Income Countries. *International Journal of Behavioral Development*, *43*(1), 74–79. https://doi.org/10.1177/0165025418783272

Last, C. G., Hansen, C., & Franco, N. (1997). Anxious children in adulthood: A prospective study of adjustment. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*(5), 645–652. https://doi.org/10.1097/00004583-199705000-00015

Lee, D., & Jackson, M. (2017). The Simultaneous Effects of Socioeconomic Disadvantage and Child Health on Children's Cognitive Development. *Demography*, *54*(5), 1845–1871. https://doi.org/10.1007/s13524-017-0605-z

Lund, C. (2014). Poverty and mental health: Towards a research agenda for low and middle-income countries. Commentary on Tampubolon and Hanandita (2014). *Social Science & Medicine*, *111*, 134–136. https://doi.org/10.1016/j.socscimed.2014.04.010

Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., Swartz, L., & Patel, V. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. *Social Science & Medicine (1982)*, *71*(3), 517–528. https://doi.org/10.1016/j.socscimed.2010.04.027

Lund, C., & Cois, A. (2018). Simultaneous social causation and social drift: Longitudinal analysis of depression and poverty in South Africa. *Journal of Affective Disorders*, *229*, 396–402. https://doi.org/10.1016/j.jad.2017.12.050

Marees, A. T., Smit, D. J. A., Abdellaoui, A., Nivard, M. G., van den Brink, W., Denys, D., Galama, T. J., Verweij, K. J. H., & Derks, E. M. (2021). Genetic correlates of socio-economic status influence the pattern of shared heritability across mental health traits. *Nature Human Behaviour*, *5*(8), Article 8. https://doi.org/10.1038/s41562-021-01053-4

Marmot, M., & Equity, U. I. of H. (2014). *Review of social determinants and the health divide in the WHO European Region: Final report*. World Health Organization. Regional Office for Europe. https://apps.who.int/iris/handle/10665/108636

Maselko, J. (2017). Social Epidemiology and Global Mental Health: Expanding the Evidence from High-Income to Low- and Middle-Income Countries. *Current Epidemiology Reports*, *4*(2), 166–173. https://doi.org/10.1007/s40471-017-0107-y

Maselko, J., Bates, L., Bhalotra, S., Gallis, J. A., O'Donnell, K., Sikander, S., & Turner, E. L. (2018). Socioeconomic status indicators and common mental disorders: Evidence from a study of prenatal depression in Pakistan. *SSM - Population Health*, *4*, 1–9. https://doi.org/10.1016/j.ssmph.2017.10.004

McGrath, J. J., Lim, C. C. W., Plana-Ripoll, O., Holtz, Y., Agerbo, E., Momen, N. C., Mortensen, P. B., Pedersen, C. B., Abdulmalik, J., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Bromet, E. J., Bruffaerts, R., Bunting, B., Almeida, J. M. C. de, Girolamo, G. de, Vries, Y. A. D., Florescu, S., … Jonge, P. de. (2020). Comorbidity within mental disorders: A comprehensive analysis based on 145 990 survey respondents from 27 countries. *Epidemiology and Psychiatric Sciences*, *29*, e153. https://doi.org/10.1017/S2045796020000633

McLaughlin, K. A., Breslau, J., Green, J. G., Lakoma, M. D., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2011). Childhood socio-economic status and the onset, persistence, and severity of DSM-IV mental disorders in a US national sample. *Social Science & Medicine*, *73*(7), 1088–1096. https://doi.org/10.1016/j.socscimed.2011.06.011

McLaughlin, K. A., Costello, E. J., Leblanc, W., Sampson, N. A., & Kessler, R. C. (2012). Socioeconomic Status and Adolescent Mental Disorders. *American Journal of Public Health*, *102*(9), 1742–1750. https://doi.org/10.2105/AJPH.2011.300477

Melchior, M., Touchette, É., Prokofyeva, E., Chollet, A., Fombonne, E., Elidemir, G., & Galéra, C. (2014). Negative events in childhood predict trajectories of internalizing symptoms up to young adulthood: An 18-year longitudinal study. *PloS One*, *9*(12), e114526. https://doi.org/10.1371/journal.pone.0114526

Middeldorp, C. M., Lamb, D. J., Vink, J. M., Bartels, M., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2014). Child care, socio-economic status and problem behavior: A study of gene-environment interaction in young Dutch twins. *Behavior Genetics*, *44*(4), 314–325. https://doi.org/10.1007/s10519-014-9660-z

Miller, P., Betancur, L., Whitfield, K., & Votruba-Drzal, E. (2021). Examining income dynamics and externalizing and internalizing trajectories through a developmental psychopathology lens: A nationally representative study. *Development and Psychopathology*, *33*(1), 1–17. https://doi.org/10.1017/S0954579419001494

Mills, M. C. (2022). Sociogenomics: Theoretical and empirical challenges of integrating molecular genetics into sociological thinking. In *Handbook of Sociological Science* (pp. 250–270). Edward Elgar Publishing. https://www.elgaronline.com/edcollchap-oa/book/9781789909432/book-part-9781789909432-22.xml

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology*, *13*(2), 355–375. https://doi.org/10.1017/s0954579401002097

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, *110*, 406–425. https://doi.org/10.1037/0033-2909.110.3.406

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2009). Risk Factors for Learning-Related Behavior Problems at 24 Months of Age: Population-Based Estimates. *Journal of Abnormal Child Psychology*, *37*(3), 401–413. https://doi.org/10.1007/s10802-008-9279-8

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. *International Journal of Epidemiology*, *47*(1), 226–235. https://doi.org/10.1093/ije/dyx206

Murray, A. L., Molenaar, D., Johnson, W., & Krueger, R. F. (2016). Dependence of Gene-by-Environment Interactions (GxE) on Scaling: Comparing the Use of Sum Scores, Transformed Sum Scores and IRT Scores for the Phenotype in Tests of GxE. *Behavior Genetics*, *46*(4), 552–572. https://doi.org/10.1007/s10519-016-9783-5

Najman, J. M., Aird, R., Bor, W., O'Callaghan, M., Williams, G. M., & Shuttlewood, G. J. (2004). The generational transmission of socioeconomic inequalities in child cognitive development and emotional health. *Social Science & Medicine*, *58*(6), 1147–1158. https://doi.org/10.1016/S0277-9536(03)00286-7

Najman, J. M., Hayatbakhsh, M. R., Clavarino, A., Bor, W., O'Callaghan, M. J., & Williams, G. M. (2010). Family Poverty Over the Early Life Course and Recurrent Adolescent and Young Adult Anxiety and Depression: A Longitudinal Study. *American Journal of Public Health*, *100*(9), 1719–1723. https://doi.org/10.2105/AJPH.2009.180943

Neale, M. C., & Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families*. Springer Netherlands. https://doi.org/10.1007/978-94-015-8018-2

Nohr, E. A., Frydenberg, M., Henriksen, T. B., & Olsen, J. (2006). Does low participation in cohort studies induce bias? *Epidemiology (Cambridge, Mass.)*, *17*(4), 413–418. https://doi.org/10.1097/01.ede.0000220549.14177.60

Nohr, E. A., & Liew, Z. (2018). How to investigate and adjust for selection bias in cohort studies. *Acta Obstetricia et Gynecologica Scandinavica*, *97*(4), 407–416. https://doi.org/10.1111/aogs.13319

Ozer, E. J., Fernald, L. C., Weber, A., Flynn, E. P., & VanderWeele, T. J. (2011). Does alleviating poverty affect mothers' depressive symptoms? A quasi-experimental investigation of Mexico's Oportunidades programme. *International Journal of Epidemiology*, *40*(6), 1565–1576. https://doi.org/10.1093/ije/dyr103

Patel, V., Burns, J. K., Dhingra, M., Tarver, L., Kohrt, B. A., & Lund, C. (2018). Income inequality and depression: A systematic review and meta-analysis of the association and a scoping review of mechanisms. *World Psychiatry*, *17*(1), 76–89. https://doi.org/10.1002/wps.20492

Pedersen, M. L., Holen, S., Lydersen, S., Martinsen, K., Neumer, S.-P., Adolfsen, F., & Sund, A. M. (2019). School functioning and internalizing problems in young schoolchildren. *BMC Psychology*, *7*(1), 88. https://doi.org/10.1186/s40359-019-0365-1

Peverill, M., Dirks, M. A., Narvaja, T., Herts, K. L., Comer, J. S., & McLaughlin, K. A. (2021). Socioeconomic status and child psychopathology in the United States: A meta-analysis of population-based studies. *Clinical Psychology Review*, *83*, 101933. https://doi.org/10.1016/j.cpr.2020.101933

Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: Theory, evidence, and policy implications. *Journal of Health and Social Behavior*, *51 Suppl*, S28-40. https://doi.org/10.1177/0022146510383498

Pine, D. S., Cohen, P., Gurley, D., Brook, J., & Ma, Y. (1998). The risk for early-adulthood anxiety and depressive disorders in adolescents with anxiety and depressive

disorders. *Archives of General Psychiatry*, *55*(1), 56–64. https://doi.org/10.1001/archpsyc.55.1.56

Plana-Ripoll, O., Pedersen, C. B., Holtz, Y., Benros, M. E., Dalsgaard, S., de Jonge, P., Fan, C. C., Degenhardt, L., Ganna, A., Greve, A. N., Gunn, J., Iburg, K. M., Kessing, L. V., Lee, B. K., Lim, C. C. W., Mors, O., Nordentoft, M., Prior, A., Roest, A. M., … McGrath, J. J. (2019). Exploring Comorbidity Within Mental Disorders Among a Danish National Population. *JAMA Psychiatry*, *76*(3), 259–270. https://doi.org/10.1001/jamapsychiatry.2018.3658

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

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), Article 7. https://doi.org/10.1038/ng.3285

Pryor, L., Strandberg-Larsen, K., Nybo Andersen, A.-M., Hulvej Rod, N., & Melchior, M. (2019). Trajectories of family poverty and children's mental health: Results from the Danish National Birth Cohort. *Social Science & Medicine*, *220*, 371–378. https://doi.org/10.1016/j.socscimed.2018.10.023

Purcell, S. (2002). Variance Components Models for Gene–Environment Interaction in Twin Analysis. *Twin Research and Human Genetics*, *5*(6), 554–571. https://doi.org/10.1375/twin.5.6.554

Reiss, D., Leve, L. D., & Neiderhiser, J. M. (2013). How genes and the social environment moderate each other. *American Journal of Public Health*, *103 Suppl 1*, S111-121. https://doi.org/10.2105/AJPH.2013.301408

Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science & Medicine (1982)*, *90*, 24–31. https://doi.org/10.1016/j.socscimed.2013.04.026

Reiss, F., Meyrose, A.-K., Otto, C., Lampert, T., Klasen, F., & Ravens-Sieberer, U. (2019). Socioeconomic status, stressful life situations and mental health problems in children and adolescents: Results of the German BELLA cohort-study. *PLOS ONE*, *14*(3), e0213700. https://doi.org/10.1371/journal.pone.0213700

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119–133. https://doi.org/10.1093/bib/3.2.119

Rimfeld, K., Krapohl, E., Trzaskowski, M., Coleman, J. R. I., Selzam, S., Dale, P. S., Esko, T., Metspalu, A., & Plomin, R. (2018). Genetic influence on social outcomes during and after the Soviet era in Estonia. *Nature Human Behaviour*, *2*(4), 269–275. https://doi.org/10.1038/s41562-018-0332-5

Rowe, D. C., Jacobson, K. C., & Van den Oord, E. J. C. G. (1999). Genetic and Environmental Influences on Vocabulary IQ: Parental Education Level as Moderator. *Child Development*, *70*(5), 1151–1162. https://doi.org/10.1111/1467-8624.00084

Rutter, M. (2006). *Genes and behavior: Nature-nurture interplay explained* (pp. viii, 280). Blackwell Publishing.

Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: Multiple varieties but real effects. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 226–261. https://doi.org/10.1111/j.1469-7610.2005.01557.x

Samuelsson, S., Byrne, B., Quain, P., Wadsworth, S., Corley, R., DeFries, J. C., Willcutt, E., & Olson, R. (2005). Environmental and genetic influences on prereading skills in Australia, Scandinavia, and the United States. *Journal of Educational Psychology*, *97*(4), 705–722. https://doi.org/10.1037/0022-0663.97.4.705

Saxena, S., Paraje, G., Sharan, P., Karam, G., & Sadana, R. (2006). The 10/90 divide in mental health research: Trends over a 10-year period. *The British Journal of Psychiatry: The Journal of Mental Science*, *188*, 81–82. https://doi.org/10.1192/bjp.bp.105.011221

Scarr-Salapatek, S. (1971). Race, Social Class, and IQ. *Science*, *174*(4016), 1285–1295. https://doi.org/10.1126/science.174.4016.1285

Selita, F., & Kovas, Y. (2019). GENES AND GINI: WHAT INEQUALITY MEANS FOR HERITABILITY. *Journal of Biosocial Science*, *51*(1), 18–47. https://doi.org/10.1017/S0021932017000645

Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene-environment interactions: Retrospect and prospect. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *60 Spec No 1*, 65–76. https://doi.org/10.1093/geronb/60.special_issue_1.65

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., … 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

South, S. C., Hamdi, N. R., & Krueger, R. F. (2017). Biometric Modeling of Gene-Environment Interplay: The Intersection of Theory and Method and Applications for Social Inequality. *Journal of Personality*, *85*(1), 22–37. https://doi.org/10.1111/jopy.12231

South, S. C., & Krueger, R. F. (2008). Marital quality moderates genetic and environmental influences on the internalizing spectrum. *Journal of Abnormal Psychology*, *117*(4), 826–837. https://doi.org/10.1037/a0013499

South, S. C., & Krueger, R. F. (2011). Genetic and environmental influences on internalizing psychopathology vary as a function of economic status. *Psychological Medicine*, *41*(1), 107–117. https://doi.org/10.1017/S0033291710000279

South, S. C., & Krueger, R. F. (2013). Marital satisfaction and physical health: Evidence for an orchid effect. *Psychological Science*, *24*(3), 373–378. https://doi.org/10.1177/0956797612453116

Strachan, E., Duncan, G., Horn, E., & Turkheimer, E. (2017). Neighborhood deprivation and depression in adult twins: Genetics and gene×environment interaction. *Psychological Medicine*, *47*(4), 627–638. https://doi.org/10.1017/S0033291716002622

Strohschein, L. (2005). Household Income Histories and Child Mental Health Trajectories. *Journal of Health and Social Behavior*, *46*(4), 359–375. https://doi.org/10.1177/002214650504600404

Taylor, A. E., Jones, H. J., Sallis, H., Euesden, J., Stergiakouli, E., Davies, N. M., Zammit, S., Lawlor, D. A., Munafò, M. R., Davey Smith, G., & Tilling, K. (2018). Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, *47*(4), 1207–1216. https://doi.org/10.1093/ije/dyy060

Torvik, F. A., Eilertsen, E. M., McAdams, T. A., Gustavson, K., Zachrisson, H. D., Brandlistuen, R., Gjerde, L. C., Havdahl, A., Stoltenberg, C., Ask, H., & Ystrom, E. (2020). Mechanisms linking parental educational attainment with child ADHD, depression,

and academic problems: A study of extended families in The Norwegian Mother, Father and Child Cohort Study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. https://doi.org/10.1111/jcpp.13197

Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., Dale, P. S., & Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence*, *42*, 83–88. https://doi.org/10.1016/j.intell.2013.11.002

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (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

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socioeconomic status: Gene-environment interaction. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(7), 734–743. https://doi.org/10.1111/j.1469-7610.2005.01552.x

Tyrrell, J., Zheng, J., Beaumont, R., Hinton, K., Richardson, T. G., Wood, A. R., Smith, G. D., Frayling, T. M., & Tilling, K. (2020). Genetic predictors of participation in optional components of UK Biobank. *BioRxiv*, 2020.02.10.941328. https://doi.org/10.1101/2020.02.10.941328

van der Sluis, S., Posthuma, D., & Dolan, C. V. (2012). A Note on False Positives and Power in G × E Modelling of Twin Data. *Behavior Genetics*, *42*(1), 170–186. https://doi.org/10.1007/s10519-011-9480-3

Vollebergh, W. A. M., van Dorsselaer, S., Monshouwer, K., Verdurmen, J., van der Ende, J., & ter Bogt, T. (2006). Mental health problems in early adolescents in the Netherlands. *Social Psychiatry and Psychiatric Epidemiology*, *41*(2), 156–163. https://doi.org/10.1007/s00127-005-0979-x

Whitfield, K., Betancur, L., Miller, P., & Votruba-Drzal, E. (2021). Longitudinal Links between Income Dynamics and Young Adult Socioeconomic and Behavioral Health Outcomes. *Youth & Society*, *53*(7), 1181–1210. https://doi.org/10.1177/0044118X21996382

WHO. (2022). *World mental health report: Transforming mental health for all* (ISBN 978-92-4-004933-8). Geneva: World Health Organisation. https://www.who.int/publications-detail-redirect/9789240049338

Wilkinson, R., Marmot, M., & Europe, W. H. O. R. O. for. (2003). *Social determinants of health: The solid facts*. World Health Organization. Regional Office for Europe. https://apps.who.int/iris/handle/10665/326568

World Health Organization. (2015). *International statistical classification of diseases and related health problems*. World Health Organization. https://apps.who.int/iris/handle/10665/246208

World Health Organization. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. http://www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/

Zavos, H. M. S., Dalton, B., Jayaweera, K., Harber-Aschan, L., Pannala, G., Adikari, A., Hatch, S. L., Siribaddana, S., Sumathipala, A., Hotopf, M., & Rijsdijk, F. V. (2020). The relationship between independent and dependent life events and depression symptoms in Sri Lanka: A twin and singleton study. *Social Psychiatry and Psychiatric Epidemiology*, *55*(2), 237–249. https://doi.org/10.1007/s00127-019-01765-z

2. Socioeconomic status and risk for child psychopathology: Exploring gene-environment interaction in the presence of gene-environment correlation using extended families in the Norwegian Mother, Father and Child Birth Cohort Study

This chapter is adapted from a manuscript that has been accepted for publication in the Journal of Child Psychology and Psychiatry. Supplementary materials for this chapter, as detailed in the text, are included in Appendix A.

Badini, I., Ahmadzadeh, Y. I., Wechsler, D. L., Lyngstad, T. H., Rayner, C., Eilersten, E. M., Zavos, H. M.S., Ystrom, E., McAdams, T. A. (*in press*). Socioeconomic status and risk for child psychopathology: Exploring gene-environment interaction in the presence of gene-environment correlation using extended families in the Norwegian Mother, Father and Child Birth Cohort Study. *Journal of Child Psychology and Psychiatry.*

2.1. Abstract

Background: Low socioeconomic status (SES) is associated with increased risk for emotional and behavioural problems among children. Evidence from twin studies has shown that family SES moderates genetic and environmental influences on child mental health. However, it is also known that SES is itself under genetic influence and previous gene-environment interaction (GxE) studies have not incorporated the potential genetic overlap between child mental health and family SES into GxE analyses. We applied a novel approach using extended family data to investigate the moderation of aetiological influences on child emotional and behavioural problems by parental socioeconomic status in the presence of modelled gene-environment correlation.

Methods: The sample comprised >28,100 children in extended-family units drawn from the Norwegian Mother, Father and Child Cohort Study (MoBa). Mothers reported children's emotional and behavioural symptoms. Parents' income and educational attainment were obtained through linkage to administrative register data. Bivariate moderation Multiple-Children-of-Twins-and-Siblings (MCoTS) models were used to analyse relationships between offspring outcomes (emotional and behavioural symptom scores) and parental socioeconomic moderators (income rank and educational attainment).

Results: The aetiology of child emotional symptoms was moderated by maternal and paternal educational attainment. Shared environmental influences on child emotional symptoms were greater at lower levels of parents' education. The aetiology of child behavioural symptoms was moderated by maternal, but not paternal, socioeconomic factors. Genetic factors shared between maternal income and child behavioural symptoms were greater in families with lower levels maternal income. Nonshared environmental influences on child behavioural symptoms were greater in families with higher maternal income and education.

Conclusions: Parental socioeconomic indicators moderated familial influences and non-shared environmental influences on child emotional and behavioural outcomes. Maternal SES and child mental health share aetiological overlap such that shared genetic influence was greater at the lower end of the socioeconomic distribution. Our findings collectively highlight

the role that family socioeconomic factors play in shaping the origins of child emotional and behavioural problems.

2.2. Introduction

Socioeconomic disadvantage is an important indicator of environmental adversity implicated in the development of mental health conditions (Costello et al., 2003; Glymour et al., 2014; Kinge et al., 2021; Maggi et al., 2010). A systematic review from 23 countries indicated that children from families with low socioeconomic status (SES) were two to three times more likely to develop emotional and behavioural problems than children from more socioeconomically advantaged families (Reiss, 2013). Low family SES is associated with disadvantages that may affect children's mental health, such as material hardship, poor-quality housing conditions, social deprivation, and exposure to stressful life situations (Braveman & Gottlieb, 2014; Wilkinson & Marmot, Michael, 2003).

When studying the effects of family socioeconomic conditions on child development, researchers use a range of indices intended to capture access to social and material resources (Diemer et al., 2013; Wilkinson & Marmot, Michael, 2003). Typically, these indices comprise one or a combination of parental educational attainment, income, and occupational status (Diemer et al., 2013). Although these indicators are moderately correlated, each has been shown to measure distinctive aspects of the socioeconomic environment (Duncan & Magnuson, 2003; Geyer et al., 2006). For example, previous studies have shown that parental income and educational attainment have stronger associations with child mental health than other measures of family SES (Lansford et al., 2019; Reiss, 2013).

Besides correlating with child mental health, the socioeconomic environment has been shown to moderate the contribution of genetic and environmental influences on child mental health (Rutter et al., 2006). Twin studies of emotional and behavioural problems in children and adolescents have reported lower heritability and higher shared environmental influences (influences that make family members more similar to one another) in low SES families compared to high SES families (Burt et al., 2016, 2020; Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). This is an example of gene-environment interaction (GxE),

whereby genetic influences on child mental health interact with the socioeconomic environment (Eaves et al., 1977). The findings from these studies could be taken as support for the bioecological framework, which proposes that more advantageous environments allow for greater expression of genetic differences, while more disadvantageous environments suppress them (Bronfenbrenner & Ceci, 1994).

As well as moderating familial influences on child mental health, family SES has also been found to moderate non-shared environmental influences (influences that make family members different from one another) on child behavioural problems (Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). Results are inconsistent on the direction of this moderation effect and depends on the indicators of family SES used (Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). This suggests that different indicators of SES may exert different moderating effects on child outcomes (Duncan & Magnuson, 2003; Geyer et al., 2006).

A shortcoming of previous twin studies that have examined the moderation of aetiological influences by family SES is that they do not model genetic overlap between SES and child psychopathology (i.e., gene-environment correlation). This has largely been due to limitations in the data used: child twin data cannot be used to estimate the heritability of environments that are entirely shared by twins, rendering it impossible to estimate the heritability of family SES (Purcell, 2002; Rijsdijk & Sham, 2002). Research using genome-wide genotype data have however shown that indicators of family SES are heritable (Krapohl & Plomin, 2016; Trzaskowski et al., 2014). Studies have also found evidence for genetic overlap between parental socioeconomic factors and child mental health (Krapohl & Plomin, 2016; Torvik et al., 2020; Trzaskowski et al., 2014). This suggests that parents are a source of genetic risk for offspring mental health and the socioeconomic conditions children grow up in, which is evidence of passive gene-environment correlation (rGE) (Kendler & Baker, 2007; Plomin et al., 1977). Genetic influence in the relationship between SES and mental health does not preclude a causal relationship between family SES and child mental health, but it does mean that genetically informed approaches are required to gain insight into the likely mechanisms underlying their association.

Previous GxE twin studies of family SES and child psychopathology have regressed out the main effect of SES on child psychopathology before testing for GxE to account for inflation of test statistics due to rGE (Purcell, 2002; van der Sluis et al., 2012). As such, these studies specifically focus on whether family SES moderates the aetiological influences *unique* to child psychopathology. This does not allow for investigation of either the nature of the covariance between family SES and child psychopathology, nor the moderation of the variance in child psychopathology that is shared with family SES. This is an issue because exclusion of the common variance between them may limit our understanding of the mechanisms through which socioeconomic disadvantage influences the aetiology of child mental health outcomes.

The Multiple-Children-of-Twins-and-Siblings (MCoTS) design provides an alternative approach to examine the moderating effect of family SES on the aetiology child emotional and behavioural outcomes. The MCoTS design involves using datasets comprising related parents and their children (i.e., extended families) to partition intergenerational associations into genetic and environmental sources of (co)variation (McAdams et al., 2018). The inclusion of multiple types of relatives means that family SES can vary within extended family units. This information allows the aetiological structure of socioeconomic indicators to be calculated and thus the genetic (and environmental) covariance between parental SES and child psychopathology. The MCoTS model can therefore be adapted to test for gene-environment interaction in the presence of gene-environment correlation.

The aim of this study was to investigate whether aetiological influences on child emotional and behavioural problems vary as a function of family SES as indexed by parental income and educational attainment. We applied moderation (GxE) MCoTS models to a large population-cohort study of twins, siblings, and half-siblings, and their children in Norway. Linked population-wide, administrative register data was used to index parental income and educational attainment.

2.3. Methods

2.3.1. Sample

Data were drawn from The Norwegian Mother and Child Cohort Study (MoBa) (Magnus et al., 2016) and from national administrative registers provided by Statistics Norway. MoBa is a population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health. Participants were recruited from all over Norway from 1999-2008. The women consented to participation in 41% of pregnancies. The cohort includes ~114,500 children, 95,200 mothers, and 75,00 fathers. Kinship between participants has been identified through linkage with pedigree and zygosity information from the Medical Birth Registry of Norway and the Norwegian Twin Registry, respectively (T. S. Nilsen et al., 2013). The current sample comprised extended family units, identified via pairs of siblings (twins, full-siblings, or half-siblings) in the parent generation. Extended family units were identified separately for sibling pairs of mothers or fathers and their children (i.e., units of maternal/paternal siblings and their children were modelled separately). Within each extended family unit, data were used from up to two parent siblings and two children per parent. Parents who do not have participating extended family members were included in analyses as nuclear family units if they had more than one child in the study. Phenotypic data were drawn from version 12 of the quality assured MoBa data files. The current study also uses national register data on parents' income and educational attainment. The Norwegian system of personal identification numbers was used to link register data with MoBa data.

2.3.2. Ethical considerations

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 currently regulated by the Norwegian Health Registry Act. The current study was approved by The Regional Committees for Medical and Health Research Ethics.

2.3.3. Measures

2.3.3.1. Outcomes: Offspring emotional and behavioural symptoms

We analysed two offspring outcomes: emotional symptom scores and behavioural symptom scores. Both were measured by maternal report when children were ages 1.5, 3, and 5 years old, using items from the Child Behaviour Checklist (CBCL) for preschool children (Achenbach, 1992). This scale consists of two subscales. The internalising sub-scale includes 13 items that measure emotional symptoms. The externalising sub-scale includes 11 items that measure behavioural symptoms. Mothers reported agreement for each item based on a three-point Likert scale: 1 = Not true; 2 = Somewhat true; 3 = Very/often true. We combined item-level scores across the three measurement waves to create composite mean scores for early-life emotional (Cronbach's alpha = 0.74) and behavioural (Cronbach's alpha = 0.84) symptoms.

2.3.3.2. Moderators: Parental socioeconomic factors

We analysed the effects of four parental socioeconomic variables as moderators of offspring outcomes. These were maternal income, maternal education, paternal income, and paternal education, which were extracted from national register data. Data on parents' total income (the sum of income from work capital gains and benefits received during the calendar year [Statistics Norway, 2023]) and educational attainment from 2000 to 2013 were included, corresponding to when child outcomes were assessed in MoBa (when children were aged 1.5, 3 and 5 years old).

At each time-point, we created a measure of income rank that indicates an individual's position in the distribution of incomes within a cohort-sex-year-group (e.g., 2005 income for females/males born in 1983). The income rank measure was scaled between 0 and 1, with higher values denoting a higher income rank within the reference group. Income rank at these three time-points was used to calculate average income rank across early childhood for each parent.

At each time-point, level of educational attainment was indexed in accordance with the Norwegian Standard Classification of Education, with values ranging from 1 (primary

education) to 8 (doctoral-level education). Education-level at these three time-points was used to calculate average education-level across early childhood for each parent.

Our decision to model the effects of these socioeconomic variables separately for mothers and fathers was guided by several factors. Initially, we observed only small to moderate correlations among the different socioeconomic indicators, ranging from .07 between maternal and paternal income to .45 between maternal and paternal education (Supplementary Table S1). Despite these moderate correlations, each indicator captures unique aspects of the socioeconomic context, and it has been argued that they should not be used interchangeably or in composite form, given their reflection of different phenomena and mechanisms contributing to social disparities in mental health (Diemer et al., 2013). Analysing these socioeconomic indicators independently may therefore provide a clearer understanding of their specific relationships with offspring mental health outcomes. Furthermore, the analysis of individual indicators facilitates more robust cross-cultural comparisons, particularly when considering findings from studies conducted in populations with diverse social stratification procedures, such as low-middle income countries (Howe et al., 2012; Maselko, 2017).

2.3.3.3. Covariates and outcome adjustments

Prior to model fitting, outcomes were regressed on the following covariates: parental age, child year of birth, number of births, and child sex. Residual scores were then log transformed to correct for positive skew and all variables were standardised prior to model fitting. Transformation of non-normal sum scores to normality improves false positive rates and reduces bias in parameter estimates (Murray et al., 2016). Moderator scores were not transformed as scaling of the predictor has minimal impact on estimates of interactions (Van Hulle & Rathouz, 2015).

2.3.4. Statistical analysis

2.3.4.1. The Multiple-Children-of-Twins-and-Siblings (MCoTS) design

The Multiple-Children-of-Twins-and-Siblings (MCoTS) model is an adapted version of the standard Children-of-Twins design (McAdams et al., 2018) that includes twins, full siblings,

and half siblings in the parent generation, and up to two children per parent in the offspring generation. Like the classical twin design (Jinks & Fulker, 1970), these models compare similarities among family members of different genetic relatedness to decompose observed variance and covariance into genetic and environmental components. Differential genetic similarity among related parents means that children of monozygotic (MZ) twins (who share 100% of their DNA) are more related to their parent's co-twin and their cousins compared to children of dizygotic (DZ) twins and full siblings (who share 50% of their segregating genes). Comparing associations between different classes of relatives in such samples (e.g., correlations between uncle/aunt and niece/nephew) allows for intergenerational transmission effects to be partitioned into passive genetic transmission, passive shared environmental transmission, and direct phenotypic components (see Table 1).

The MCoTS model decomposes variance in the parent trait (e.g., SES indicators) into parent genetic (A1), shared environmental (C1), and nonshared environmental (E1) components, and variance in child traits (e.g., emotional, and behavioural problems) into child genetic (A2), shared environmental (C2), and nonshared environmental (E2) components (Figure S1). Variance explained by A1, C1 and E1 are unique to the parent generation and non-overlapping with variance explained by A2, C2 and E2 in the child generation. The intergenerational association between the parental and child trait is partitioned into genetic transmission (A1'), shared environmental effects (C1'; indexing environmental influences shared across the extended family) and residual phenotypic transmission (p; accounting for non-genetic effects shared between the parent and child, and effects of any sources of confounding unaccounted for; Figure S1).

2.3.4.2. Using the MCoTS design to investigate moderation of offspring outcomes by parental socioeconomic factors

To investigate whether parental socioeconomic factors moderate the aetiology of child emotional and behavioural problems, we adapted the MCoTS model to include moderation terms on the intergenerational and child trait paths (Figure 1). This is based on the bivariate moderation model proposed by Purcell (2002) (Supplementary Figure S2), in which it is possible to simultaneously model: 1) shared genetic and environmental effects between the moderator and trait; 2) the moderation of the genetic and environmental variance

components shared between the moderator and trait; 3) the moderation of the variance components unique to the trait. The bivariate moderation model requires the moderator variable to vary within extended families, hence it has not previously been applied to studies of twin children who share the same nuclear family environment. Because SES varies within extended family units (i.e., across adult siblings who have children), the MCoTS design allows us to utilise the bivariate moderation model to estimate whether parental SES moderates the intergenerational and unique genetic and environmental influences on child psychopathology (Figure 1).

We conducted eight bivariate moderation MCoTS models to analyse the relationships between two offspring outcomes (emotional and behavioural symptom scores) and four parental socioeconomic moderators (income rank and educational attainment). To test for the significance of moderation effects, we compared each full moderation model with a no-moderation model, in which all the moderation parameters were dropped. Significance of individual variance components was indicated by 95% confidence intervals (CI) around the moderated parameters from the full model[2]. Models were fitted using full-information-maximum-likelihood and compared using the x^2 distribution of the -2 log-likelihood model fit index and Akaike's Information Criterion (AIC) (Akaike, 1987). Analyses were performed in R version 4.0.3 using the open source package OpenMx v.2.12.1 (Neale et al., 2016).

2.4. Results

2.4.1. Descriptive statistics

Table 1 presents an overview of the study sample. Mothers at recruitment were aged on average 30.22 (SD=4.18) years and fathers 32.04 (SD = 4.88). 51% of the children included were males. Phenotypic correlations between all study variables are presented in Supplementary Table S1. Child emotional and behavioural symptom scores were moderately correlated (*r* = 0.39). Parental socioeconomic factors were negatively correlated with child

[2] We did not fit constrained sub models to test for the significance of individual moderated parameters because omission of moderation effects by fixing them to 0 can bias estimation of parameters (e.g., dropping ßc can inflate ßa, and there are issues of specificity in distinguishing between ßa and ßc; see Figure S2) (Van Hulle & Rathouz, 2015).

emotional and behavioural scores (*r* ranged from -0.04 to -0.11). Transformed variables had approximately normal distributions (Supplementary Figure S3).

2.4.2. MCoTS moderation models

2.4.2.1. Moderation of offspring mental health outcomes by maternal socioeconomic factors

Overall model fit

Four moderation MCoTS models were applied to test whether the aetiology of child emotional and behavioural symptoms were moderated by two maternal exposures: maternal income and education attainment. Dropping all moderation parameters resulted in a significant worsening in model fit compared to the full moderation models (Table S2). Moderated parameter estimates from the full moderation model are presented in Supplementary Table S3. Figures 3 and 4 show the unstandardised variance components for child emotional and behavioural problems as a function of maternal income and educational attainment.

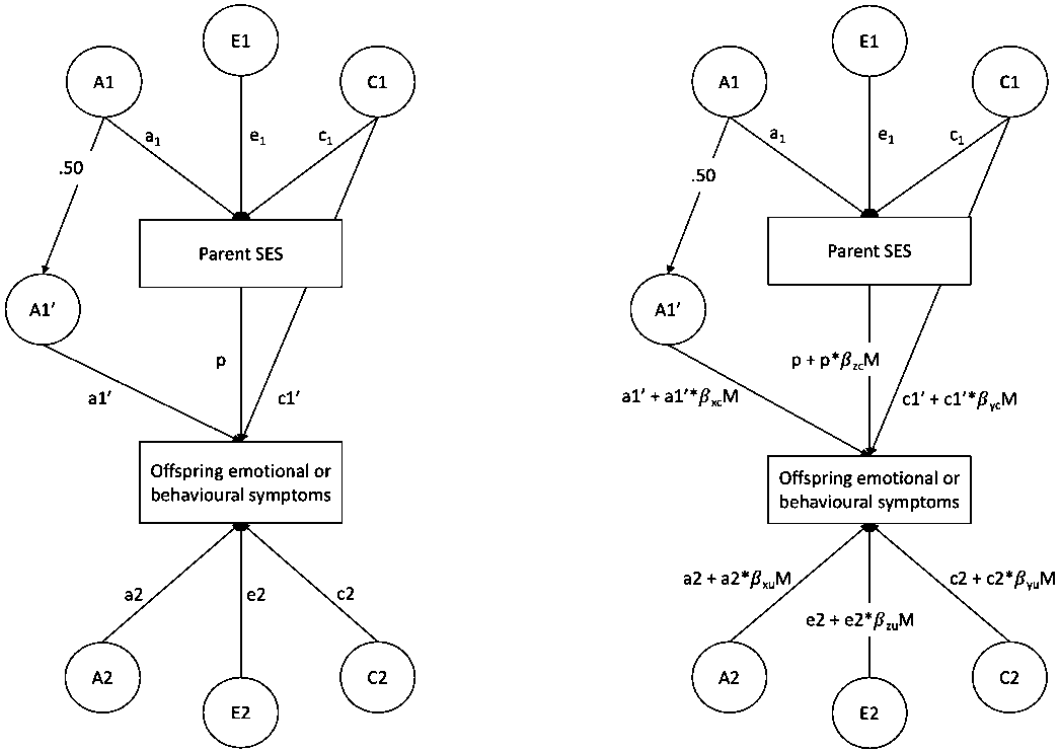


Figure 1. Partial path diagram of the Multiple-Children-of-Twins-and-Siblings (MCoTS) model showing the addition of moderation terms to the intergenerational and child trait paths (right) to the MCoTS model (left; Supplementary Figure S1). The model is shown for only one sibling in the parent generation and one child. *Right figure.* A1 = additive genetic effects on parental trait; C1 = shared environmental effects on parental trait; E1 = nonshared environmental effects on parental trait; A1' = genetic effects shared between parental trait and offspring trait; C1' = extended family effects (i.e. shared environment of the parents influences offspring trait); A2 = genetic effects specific to offspring trait; C2 = shared environmental effects on offspring trait; E2 = nonshared environmental effects on offspring trait; p = residual phenotypic association after accounting for genetic and environmental overlap. *Left figure.* A1', C1', and p are the variance components common to parent SES (the moderator) and child emotional or behavioural symptoms. A2, C2, and E2 are the variance components unique to child emotional or behavioural symptoms. β coefficients index the direction and magnitude of moderation. The total variance of the trait can be calculated by squaring and summing all the paths leading to it: $Var(T|M) = (a1' + a1'*\beta_{xc}M)^2 + (p + p*\beta_{zc}M)^2 + (c1' + c1'*\beta_{yc}M)^2 + (a2 + a2*\beta_{xu}M)^2 + (e2 + e2*\beta_{zu}M)^2 + (c2 + c2*\beta_{yu}M)^2$.

Note. Var = variance; T = trait; M = moderator; The loadings of the cross-paths connecting M to T consist of parts unrelated to the moderator M, i.e., a1', p, and c1' and parts that depend on M via weights β_{xc}, β_{zc}, and β_{yc}. The loadings of the paths unique to T consist of parts that are unrelated to M, i.e., a2, e2 and c2, and parts that depend on M via weights β_{xu}, β_{zu}, and β_{yu}.

Table 1. Study sample size stratified by maternal/paternal kinship

Mother-child extended families (N = 16408)		
Study sample size stratified by mothers' relatedness	**rA**	**n**
Identical twin pair	1.0	57
Full-sibling/fraternal twin pair	.50	4963
Half-sibling pair	.25	362
Unrelated (sibling-in-law) pair	.00	11026
Number of offspring pairs linked to each mother	**rA**	**n**
Full-sibling pair	.50	4346
Maternal half-sibling pair	.25	32
Unpaired (single) offspring	---	16692
Number of offspring pairs linked to each mother in unpaired nuclear families	**rA**	**n**
Identical twin pair	1.0	167
Full-sibling/fraternal twin pair	.50	6859
Maternal half-sibling pair	.25	56
Father-child extended families (N = 16455)		
Study sample size stratified by fathers' relatedness	**rA**	**n**
Identical twin pair	1.0	23
Full-sibling/fraternal twin pair	.50	3464
Half-sibling pair	.25	164
Unrelated (sibling-in-law) pair	.00	12804
Number of offspring pairs linked to each father	**rA**	**n**
Full-sibling pair	.50	4299
Paternal half-sibling pair	.25	16
Unpaired (single) offspring	---	16845
Number of offspring pairs linked to each father in unpaired nuclear families	**rA**	**n**
Identical twin pair	1.0	167
Full-sibling/fraternal twin pair	.50	6859
Paternal half-sibling pair	.25	0

Emotional symptoms

Lower levels of maternal income and education were associated with greater variation in child emotional symptoms (Figure 2a and Figure 2b, respectively). Figure 2a suggests genetic influences on child emotional symptoms (A1' and A2 variance components) and shared environmental influences unique to child emotional outcomes (C2 variance component) were greater at lower versus higher levels of maternal income. Although moderation terms could not all be dropped from the model without a significant loss of model fit (Table S2), no single variance component for emotional outcomes appeared significantly moderated by maternal income when confidence intervals around the moderation terms were inspected (see

Supplementary Table S3). Figure 2b suggests moderation effects on the A1' and C2 variance components, such that genetic factors common to maternal education and child emotional outcomes (A1'), and shared environmental influences unique to child emotional outcomes (C2), were greater at lower levels of maternal education. Examination of the confidence intervals around the moderation parameter estimates from the full model showed significant moderation on the shared environmental component unique to child emotional outcomes (C2) as a function of maternal education (β_{yu} = -0.06, 95% CI [-0.10, -0.01]; Supplementary Table S3).

Behavioural symptoms

Total phenotypic variation in child behavioural symptoms was stable across levels of maternal income (Figure 3a) and education (Figure 3b). Figure's 3a and 3b suggest however that genetic influences common to maternal socioeconomic factors and child behavioural outcomes (A1') were greater at lower levels of maternal income and education, whereas nonshared environmental factors (E2) showed a stronger influence on child behavioural outcomes with increasing maternal income and education. Examination of the confidence intervals around the moderation terms from the full model showed significant moderation on the shared genetic component (A1'; β_{xc} = 0.06, 95% CI [0.01, 0.09]) and the nonshared environmental component unique to child emotional outcomes (E2; β_{zu} = 0.06, 95% CI [0.02, 0.09]) as a function of maternal income (Supplementary Table S3). Significant moderation was observed on the nonshared environmental component (E2) as a function of maternal education (β_{zu} = 0.06, 95% CI [0.02, 0.08]; Supplementary Table S3).

2.4.2.2. Moderation of offspring mental health outcomes by paternal socioeconomic factors

Overall model fit

Four moderation MCoTS models were applied to test whether the aetiology of child emotional and behavioural symptoms were moderated by two paternal exposures: income and educational attainment. Dropping all moderation parameters resulted in a significant decrease in fit compared to the full models for child emotional and behavioural symptoms

(Table S4). Moderated parameter estimates from the full moderation model are presented in Supplementary Table S5.

Emotional symptoms

Total phenotypic variance in child emotional symptoms was greater at low versus high levels of paternal income (Figure 4a) and education (Figure 4b). Figure 4a suggests that shared (C2) and non-shared environmental influences (E2) unique to child emotional outcomes were greater at lower income levels, whereas genetic influences unique to child emotional outcomes (A2) increased with increasing paternal income. Although moderation terms could not all be dropped from the model without a significant loss of model fit (Table S3), confidence intervals around the moderation terms did not highlight any single variance component as significantly moderated by paternal income (Supplementary Table S5). Figure 4b suggests moderation of the shared environmental variance unique to child emotional outcomes (C2), which was greater at lower versus higher levels of paternal education. Examination of the confidence intervals around the moderation parameter estimates from the full model showed significant moderation on the shared environmental component unique to child emotional outcomes (C2) as a function of paternal education (β_{yu} = -0.05, 95% CI [-0.07, -0.02]; Supplementary Table S5).

Behavioural symptoms

No significant moderation effects of paternal socioeconomic factors were found for child behavioural symptoms (Table S4).

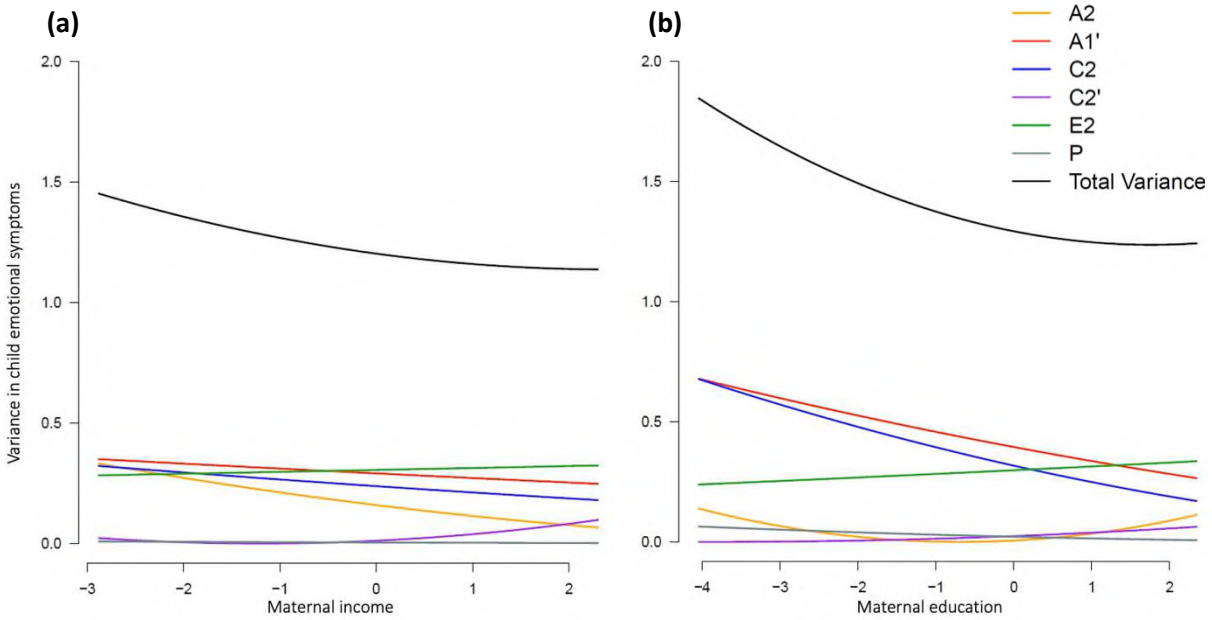


Figure 2. Unstandardised variance components in child emotional problems moderated by maternal income rank (a) and educational attainment (b)

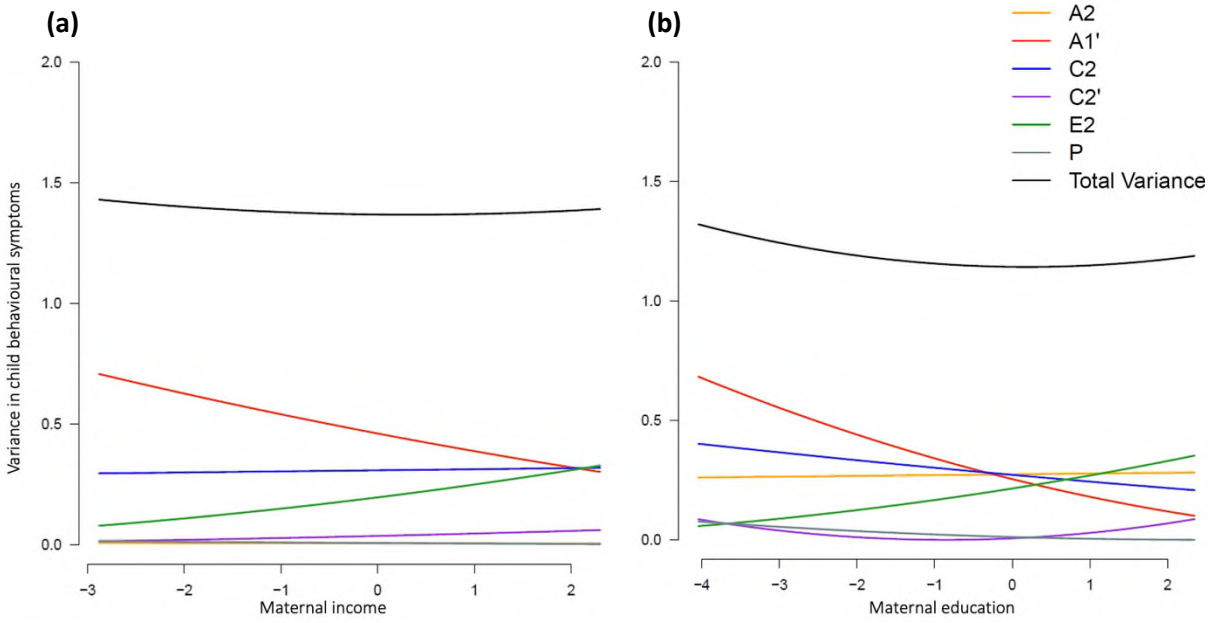


Figure 3. Unstandardised variance components in child behavioural problems moderated by maternal income rank (a) and education attainment (b)

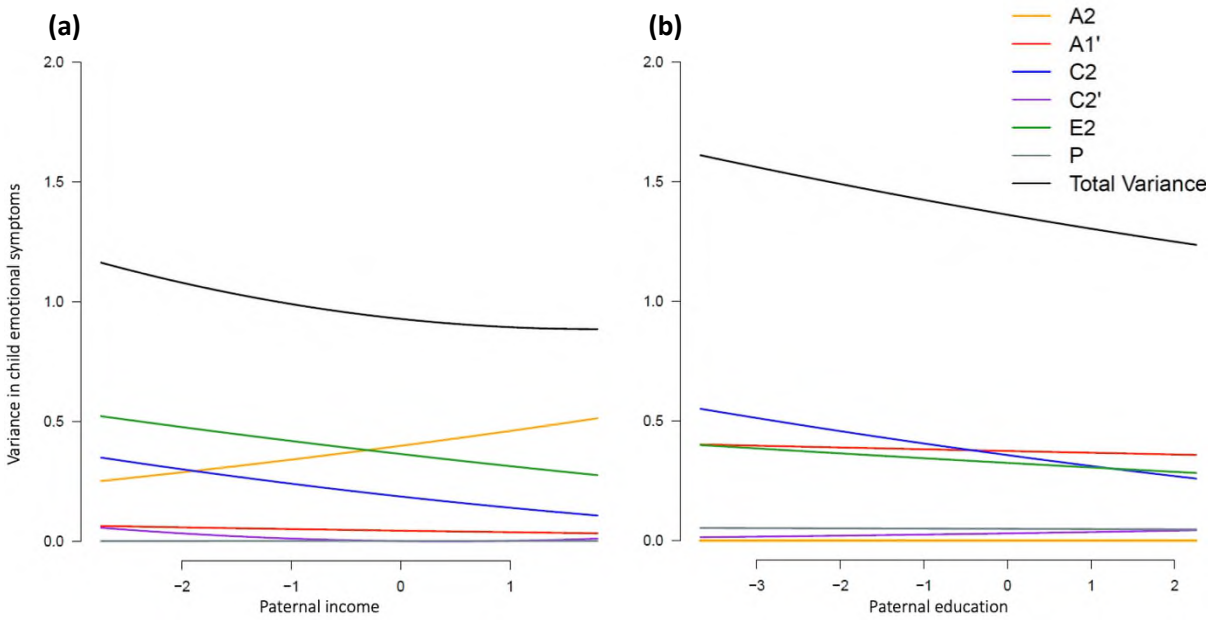


Figure 4. Unstandardised variance components in child emotional problems moderated by paternal income rank (a) and educational attainment (b).

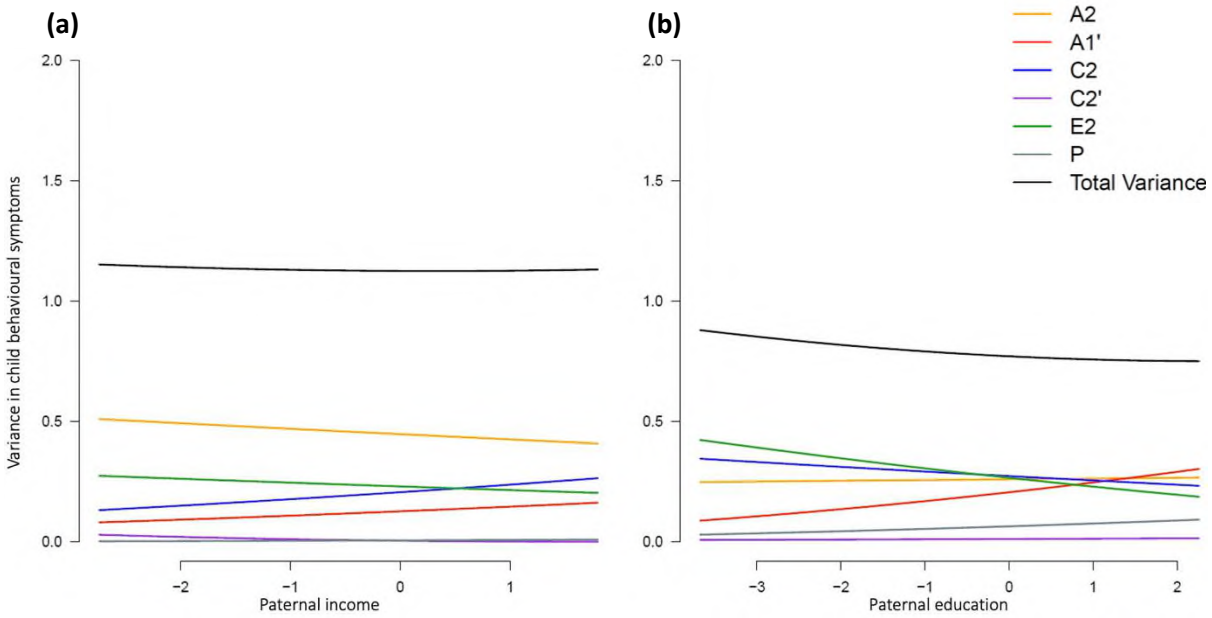


Figure 5. Unstandardised variance components in child behavioural problems moderated by paternal income rank (a) and educational attainment (b).

2.5. Discussion

We examined whether parental socioeconomic factors significantly moderated aetiological influences on child emotional and behavioural symptoms (GxE). In contrast to previous studies, we were able to model gene-environment correlation (rGE) and GxE simultaneously. Using extended family units, we were able to estimate the aetiological structure of socioeconomic factors and thus the genetic and environmental covariance between parents' socioeconomic factors and child outcomes. This allowed us to investigate whether family SES moderates the influence of genetic and environmental influences unique to child emotional and behavioural problems, as well as any genetic and environmental influences shared with parental socioeconomic factors.

We found evidence that shared environmental influences unique to child emotional symptoms (C2) were higher at lower levels of maternal and paternal educational attainment. Thus, shared environmental influences explained greater variance in emotional symptoms for children in more socioeconomically disadvantaged circumstances. This is consistent with previous evidence indicating that shared environmental influences on child emotional problems have a greater influence among children from families with lower SES (Middeldorp et al., 2014). In contrast to prior studies, we did not observe significant moderation on shared environmental influences specific to child behavioural symptoms (Burt et al., 2016, 2020; Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). Discrepancies could in part reflect the different measures used to index family SES. In this study, we used individual-level national register data to obtain independent information on both maternal and paternal income levels and educational attainment. In contrast, past research used either measures of neighbourhood-level disadvantage (Burt et al., 2016, 2020) or self-report data on parents' educational attainment and occupational status, and did not distinguish paternal from maternal socioeconomic indicators (Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). Reporting and response biases have been shown to affect the validity of self-report measures of SES, leading to difficulties with comparisons between studies (Angel et al., 2019; Duncan & Magnuson, 2003; Lorant et al., 2007; Moore et al., 2000). Differences between study samples, such as sample age (e.g., child outcomes assessed in early childhood versus middle childhood/adolescence) and that samples were drawn from different populations

(e.g., Norway versus the USA), could also contribute to differences in the pattern of interaction effects detected.

We also found evidence that genetic influences on child behavioural symptoms that are shared with maternal income (A1') were greater at lower levels of income. Thus, these shared genetic factors explained more of the variance in child behavioural symptoms for families at this end of the socioeconomic distribution. This finding appears to support the diathesis stress framework (Monroe & Simons, 1991), which suggests that genetic differences in behavioural outcomes manifest in poorer environments. In turn, the influence of the non-shared environment on child behavioural outcomes increased with increasing maternal income and educational attainment. This is consistent with previous evidence demonstrating that non-shared influences explained more of the variance in behavioural problems in children from families with higher parental educational attainment (Hendriks et al., 2019). Our finding that shared genetic factors explained greater variation in behavioural symptoms for children in lower SES families is in contrast with those of prior twin studies that reported lower heritability of behavioural problems for children in less advantaged environments (Burt et al., 2016; Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006). It is possible that previous studies have not detected increased genetic influences on child psychopathology at lower ends of the socioeconomic distribution, because the genetic effects are shared between parent SES and child psychopathology. Previous studies have regressed out this covariance and only tested for GxE on the variance that is unique to the child trait (Burt et al., 2016, 2020; Middeldorp et al., 2014; Purcell, 2002; Tuvblad et al., 2006). Our findings suggest that this may have biased findings in previous studies. In the current study, our ability to model the covariance between socioeconomic factors and child outcomes revealed that heritability was *higher* at lower levels of maternal income/education. This increased heritability was driven by genetic variance shared between maternal SES indicators and child behavioural outcomes that would have been regressed out of previous analyses. To our knowledge, ours is the first GxE study to incorporate moderation of the genetic variance that is shared between family SES and child mental health.

A possible interpretation of the finding that aetiological influences on child behavioural outcomes were moderated by maternal, but not paternal, socioeconomic indicators is that

these indicators could partly be reflective of individual traits of the parents and may affect child behaviour via different pathways. The correlation between maternal and paternal income was low (r = .07) and given that mothers still tend to take on the majority of child caregiving duties (Astrid, 2020; Dietrichson, 2017), low income for many mothers in this sample may indicate someone who is taking on the majority of the childcare but may not be materially poor. Future work disentangling the respective contributions of different socioeconomic indicators would improve our understanding of the mechanisms underlying intergenerational associations between family SES and child mental health. While we modelled mothers and fathers separately, future analyses could also seek to model them simultaneously as an overall measure of family SES and investigate whether moderation effects differ between individual and composite measures of SES.

2.5.1. Limitations

There are some limitations to our study. First, income inequality in Norway is low relative to many other countries (UNICEF, 2017), so findings from this sample may not generalise to countries with greater economic disparity. Second, participation in MoBa is characterised by self-selection. MoBa participants have been found to have a higher educational attainment and experience lower levels of mental health problems compared to those who did not participate, which may affect generalisability of our findings (R. M. Nilsen et al., 2009) However, studies suggest that reduced prevalence rates in MoBa do not necessarily lead to biases in estimates of associations between exposures and outcomes (R. M. Nilsen et al., 2009; Oerbeck et al., 2017). Third, the current results should be considered specific to early childhood. Time-specific associations between family SES and trajectories of emotional and behavioural problems have been reported (Miller et al., 2021). In addition, developmental differences in moderation could be expected given that the aetiology of mental health problems changes across development (Hannigan et al., 2017). Last, although the phenotypic associations between parental socioeconomic factors and child emotional and behavioural outcomes were in the expected direction, they were very small (r from -0.04 to -0.11). However, estimates were in line with those reported in previous studies (Burt et al., 2016, 2020; Tuvblad et al., 2006) and the presence of small phenotypic associations between the

moderator and outcome has no bearing on the extent of aetiologic moderation (van der Sluis et al., 2012).

2.6. Conclusion

This study uses a novel genetically informative research design in a large population-based sample. We test for gene-environment interaction (GxE) in the presence of gene-environment correlation (rGE) for environmental moderators that are necessarily shared between children growing up in the same family. We provide evidence that family socioeconomic factors moderate the influence of familial and non-shared environmental influences on child emotional and behaviour problems. This is the first study to demonstrate moderation of genetic variance that is shared between family SES and child mental health. Our findings indicate that the presence and pattern of moderation effects varies depending on the measure used to index family SES. Future studies may next consider using similar large-scale, genetically informative data to explore how the dynamics of different measures of the family socioeconomic environment relate to *trajectories* of aetiologic moderation on emotional and behavioural problems across child developmental periods.

2.7. Acknowledgements

During the writing of this manuscript, IB was supported by the UK Economic and Social Research Council (ESRC) and King's College London member of the ESRC Doctoral Training Partnership in Interdisciplinary Social Science (LISS DTP) (grant ST11872). DLW was supported by the UK Medical Research Council (grant MR/N013700/1) and King's College London member of the MRC Doctoral Training Partnership in Biomedical Sciences. TAM and YIA were funded by a Sir Henry Dale Fellowship awarded to TAM, jointly funded by the Wellcome Trust and the Royal Society (grant 107706/Z/15/Z). TAM, YIA and CR were funded by a Senior Research Fellowship awarded to TAM funded by the Wellcome Trust (grant number 220382/Z/20/Z). The Research Council of Norway supported EY and TAM (grant number 288083) and EME (grant 262177). The Norwegian Mother, Father and Child Cohort Study data collection was supported by the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research, NIH/NIEHS (contract N01-ES-75558), NIH/NINDS (grants 1 UO1 NS 047537-01, 2 UO1 NS 047537-06A1). The funders of the study had no role in study

design, data collection, data analysis, data interpretation, or writing of the report. We are grateful to all the participating families in Norway who take part in this on-going cohort study.

2.8. References

Achenbach, T. (1992). Manual for The Child Behavior Checklist/4-18 and 1991 Profile. *University of Vermont, Department of Psychiatry*. https://ci.nii.ac.jp/naid/20001666977/

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*(3), 317–332. https://doi.org/10.1007/BF02294359

Angel, S., Disslbacher, F., Humer, S., & Schnetzer, M. (2019). What did you really earn last year?: Explaining measurement error in survey income data. *Journal of the Royal Statistical Society Series A*, *182*(4), 1411–1437.

Astrid, K. (2020, March 16). Can women have it all? *Policy Forum*. https://www.policyforum.net/can-women-have-it-all/

Braveman, P., & Gottlieb, L. (2014). The Social Determinants of Health: It's Time to Consider the Causes of the Causes. *Public Health Reports*, *129*(Suppl 2), 19–31.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, *101*(4), 568–586. https://doi.org/10.1037/0033-295x.101.4.568

Burt, S. A., Klump, K. L., Gorman-Smith, D., & Neiderhiser, J. M. (2016). Neighborhood Disadvantage Alters the Origins of Children's Nonaggressive Conduct Problems. *Clinical Psychological Science : A Journal of the Association for Psychological Science*, *4*(3), 511–526. https://doi.org/10.1177/2167702615618164

Burt, S. A., Pearson, A. L., Carroll, S., Klump, K. L., & Neiderhiser, J. M. (2020). Child Antisocial Behavior Is more Environmental in Origin in Disadvantaged Neighborhoods: Evidence Across Residents' Perceptions and Geographic Scales in Two Samples. *Journal of Abnormal Child Psychology*, *48*(2), 265–276. https://doi.org/10.1007/s10802-019-00587-6

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships Between Poverty and PsychopathologyA Natural Experiment. *JAMA*, *290*(15), 2023–2029. https://doi.org/10.1001/jama.290.15.2023

Diemer, M. A., Mistry, R. S., Wadsworth, M. E., López, I., & Reimers, F. (2013). Best Practices in Conceptualizing and Measuring Social Class in Psychological Research. *Analyses of Social Issues and Public Policy*, *13*(1), 77–113. https://doi.org/10.1111/asap.12001

Dietrichson, S. (2017). *Mum takes more responsibility for the children, even in gender equal couples*. Kilden. https://kjonnsforskning.no/en/2017/10/mum-takes-more-responsibility-children-even-gender-equal-couples

Duncan, G. J., & Magnuson, K. A. (2003). Off with Hollingshead: Socioeconomic resources, parenting, and child development. In *Socioeconomic status, parenting, and child development* (pp. 83–106). Lawrence Erlbaum Associates Publishers.

Eaves, L. J., Last, K., Martin, N. G., & Jinks, J. L. (1977). A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. *British Journal of Mathematical and Statistical Psychology*, *30*(1), 1–42. https://doi.org/10.1111/j.2044-8317.1977.tb00722.x

Geyer, S., Hemström, Ö., Peter, R., & Vågerö, D. (2006). Education, income, and occupational class cannot be used interchangeably in social epidemiology. Empirical evidence against a common practice. *Journal of Epidemiology & Community Health*, *60*(9), 804–810. https://doi.org/10.1136/jech.2005.041319

Glymour, M. M., Avendano, M., & Kawachi, I. (2014). Socioeconomic Status and Health. In *Social Epidemiology*. Oxford University Press. https://oxfordmedicine.com/view/10.1093/med/9780195377903.001.0001/med-9780195377903-chapter-2

Hannigan, L. J., McAdams, T. A., & Eley, T. C. (2017). Developmental change in the association between adolescent depressive symptoms and the home environment: Results from a longitudinal, genetically informative investigation. *Journal of Child Psychology and Psychiatry*, *58*(7), 787–797. https://doi.org/10.1111/jcpp.12689

Hendriks, A. M., Finkenauer, C., Nivard, M. G., Van Beijsterveldt, C. E. M., Plomin, R. J., Boomsma, D. I., & Bartels, M. (2019). Comparing the genetic architecture of childhood behavioral problems across socioeconomic strata in the Netherlands and the United Kingdom. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-019-01357-x

Jinks, J. L., & Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of the human behavior. *Psychological Bulletin*, *73*(5), 311–349. https://doi.org/10.1037/h0029135

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

Kinge, J. M., Øverland, S., Flatø, M., Dieleman, J., Røgeberg, O., Magnus, M. C., Evensen, M., Tesli, M., Skrondal, A., Stoltenberg, C., Vollset, S. E., Håberg, S., & Torvik, F. A. (2021). Parental income and mental disorders in children and adolescents: Prospective register-based study. *International Journal of Epidemiology*, *dyab066*. https://doi.org/10.1093/ije/dyab066

Krapohl, E., & Plomin, R. (2016). Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Molecular Psychiatry*, *21*(3), 437–443. https://doi.org/10.1038/mp.2015.2

Lansford, J. E., Malone, P. S., Tapanya, S., Tirado, L. M. U., Zelli, A., Alampay, L. P., Al-Hassan, S. M., Bacchini, D., Bornstein, M. H., Chang, L., Deater-Deckard, K., Di Giunta, L., Dodge, K. A., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., & Steinberg, L. (2019). Household Income Predicts Trajectories of Child Internalizing and Externalizing Behavior in High-, Middle-, and Low-Income Countries. *International Journal of Behavioral Development*, *43*(1), 74–79. https://doi.org/10.1177/0165025418783272

Lorant, V., Croux, C., Weich, S., Deliège, D., Mackenbach, J., & Ansseau, M. (2007). Depression and socio-economic risk factors: 7-year longitudinal population study. *The British Journal of Psychiatry: The Journal of Mental Science*, *190*, 293–298. https://doi.org/10.1192/bjp.bp.105.020040

Maggi, S., Irwin, L. J., Siddiqi, A., & Hertzman, C. (2010). The social determinants of early child development: An overview. *Journal of Paediatrics and Child Health*, *46*(11), 627–635. https://doi.org/10.1111/j.1440-1754.2010.01817.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

McAdams, T. A., Hannigan, L. J., Eilertsen, E. M., Gjerde, L. C., Ystrom, E., & Rijsdijk, F. V. (2018). Revisiting the Children-of-Twins Design: Improving Existing Models for the Exploration of Intergenerational Associations. *Behavior Genetics*, *48*(5), 397–412. https://doi.org/10.1007/s10519-018-9912-4

Middeldorp, C. M., Lamb, D. J., Vink, J. M., Bartels, M., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2014). Child care, socio-economic status and problem behavior: A study of gene-environment interaction in young Dutch twins. *Behavior Genetics*, *44*(4), 314–325. https://doi.org/10.1007/s10519-014-9660-z

Miller, P., Betancur, L., Whitfield, K., & Votruba-Drzal, E. (2021). Examining income dynamics and externalizing and internalizing trajectories through a developmental psychopathology lens: A nationally representative study. *Development and Psychopathology*, *33*(1), 1–17. https://doi.org/10.1017/S0954579419001494

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, *110*, 406–425. https://doi.org/10.1037/0033-2909.110.3.406

Moore, J. C., Stinson, L. L., & Welniak, E. J. (2000). Income Measurement Error in Surveys: A Review. *Journal of Official Statistics*, *16*, 31.

Murray, A. L., Molenaar, D., Johnson, W., & Krueger, R. F. (2016). Dependence of Gene-by-Environment Interactions (GxE) on Scaling: Comparing the Use of Sum Scores, Transformed Sum Scores and IRT Scores for the Phenotype in Tests of GxE. *Behavior Genetics*, *46*(4), 552–572. https://doi.org/10.1007/s10519-016-9783-5

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

Nilsen, R. M., Vollset, S. E., Gjessing, H. K., Skjærven, R., Melve, K. K., Schreuder, P., Alsaker, E. R., Haug, K., Daltveit, A. K., & Magnus, P. (2009). Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatric and Perinatal Epidemiology*, *23*(6), 597–608. https://doi.org/10.1111/j.1365-3016.2009.01062.x

Nilsen, T. S., Knudsen, G. P., Gervin, K., Brandt, I., Røysamb, E., Tambs, K., Orstavik, R., Lyle, R., Reichborn-Kjennerud, T., Magnus, P., & Harris, J. R. (2013). The Norwegian Twin Registry from a public health perspective: A research update. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, *16*(1), 285–295. https://doi.org/10.1017/thg.2012.117

Oerbeck, B., Overgaard, K. R., Aspenes, S. T., Pripp, A. H., Mordre, M., Aase, H., Reichborn-Kjennerud, T., & Zeiner, P. (2017). ADHD, comorbid disorders and psychosocial functioning: How representative is a child cohort study? Findings from a national patient registry. *BMC Psychiatry*, *17*(1), 23. https://doi.org/10.1186/s12888-017-1204-7

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*(2), 309–322. https://doi.org/10.1037/0033-2909.84.2.309

Purcell, S. (2002). Variance Components Models for Gene–Environment Interaction in Twin Analysis. *Twin Research and Human Genetics*, *5*(6), 554–571. https://doi.org/10.1375/twin.5.6.554

Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. *Social Science & Medicine (1982)*, *90*, 24–31. https://doi.org/10.1016/j.socscimed.2013.04.026

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119–133. https://doi.org/10.1093/bib/3.2.119

Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: Multiple varieties but real effects. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 226–261. https://doi.org/10.1111/j.1469-7610.2005.01557.x

Statistics Norway. (2023, December 19). *Income and wealth statistics for households*. SSB. https://www.ssb.no/en/inntekt-og-forbruk/inntekt-og-formue/statistikk/inntekts-og-formuesstatistikk-for-husholdninger

Torvik, F. A., Eilertsen, E. M., McAdams, T. A., Gustavson, K., Zachrisson, H. D., Brandlistuen, R., Gjerde, L. C., Havdahl, A., Stoltenberg, C., Ask, H., & Ystrom, E. (2020).

Mechanisms linking parental educational attainment with child ADHD, depression, and academic problems: A study of extended families in The Norwegian Mother, Father and Child Cohort Study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. https://doi.org/10.1111/jcpp.13197

Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., Dale, P. S., & Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence*, *42*, 83–88. https://doi.org/10.1016/j.intell.2013.11.002

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socioeconomic status: Gene-environment interaction. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(7), 734–743. https://doi.org/10.1111/j.1469-7610.2005.01552.x

UNICEF. (2017). *Building the Future: Children and the Sustainable Development Goals in Rich Countries*. https://www.unicef-irc.org/publications/890-building-the-future-children-and-the-sustainable-development-goals-in-rich-countries.html

van der Sluis, S., Posthuma, D., & Dolan, C. V. (2012). A Note on False Positives and Power in G × E Modelling of Twin Data. *Behavior Genetics*, *42*(1), 170–186. https://doi.org/10.1007/s10519-011-9480-3

Van Hulle, C. A., & Rathouz, P. J. (2015). Operating Characteristics of Statistical Methods for Detecting Gene-by-Measured Environment Interaction in the Presence of Gene-Environment Correlation under Violations of Distributional Assumptions. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, *18*(1), 19–27. https://doi.org/10.1017/thg.2014.81

Wilkinson, R., & Marmot, Michael (Eds.). (2003). *The solid facts: Social determinants of health* (2nd ed). WHO Regional Office for Europe.

3. Associations between socioeconomic factors and depression in Sri Lanka: The role of gene-environment interplay

This chapter is adapted from a manuscript that has been accepted for publication in the Journal of Affective disorders. Supplementary materials for this chapter, as detailed in the text, are included in Appendix B.

Badini, I., Jayaweera, K., Pannala, G., Adikari, A., Siribaddana, S., Sumathipala, A., McAdams, T. A., Harber-Aschan, L., Hotopf, M., Rijsdijk, F. V., & Zavos, H. M. S. (2023). Associations between socioeconomic factors and depression in Sri Lanka: The role of gene-environment interplay. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2023.07.084

3.1. Abstract

Background: Low socioeconomic status is a risk factor for depression. The nature and magnitude of associations can differ cross-culturally and is influenced by a range of contextual factors. We examined the aetiology of socioeconomic indicators and depression symptoms and investigated whether socioeconomic indicators moderate genetic and environmental influences on depression symptoms in a Sri Lankan population.

Methods: Data were from a population-based sample of twins (N = 2934) and singletons (N = 1035) in Colombo, Sri Lanka. Standard of living, educational attainment, and financial strain were used to index socioeconomic status. Depression symptoms were assessed using the Revised Beck Depression Inventory. Structural equation modelling explored genetic and environmental influences on socioeconomic indicators and depression symptoms and moderation of aetiological influences on depression symptoms by socioeconomic status.

Results: Depression symptoms were associated with lower standard of living, lower educational attainment, and financial strain. Sex differences were evident in the aetiology of standard of living, with a small contribution of genetic influences in females. Educational attainment was moderately heritable in both males and females. Total variance in depression was greater among less socioeconomically advantaged individuals. Modest evidence of moderation of the aetiology of depression by standard of living and education was observed.

Limitations: While the sample is representative of individuals living in Colombo District, it may not be representative of different regions of Sri Lanka.

Conclusions: The aetiology of depression varies across socioeconomic contexts, suggesting a potential mechanism through which socioeconomic disadvantage increases the risk for depression in Sri Lanka. Findings have implications for cross-cultural investigations of the role of socioeconomic factors in depression and for identifying targets for social interventions.

3.2. Introduction

Major depression is highly prevalent and a leading cause of global disability (World Health Organization, 2017). The rising burden of depression worldwide disproportionately affects low- and middle-income countries (LMIC), where more than 80% of this disease burden is among people living in these countries (World Health Organization, 2017). Socioeconomic status (SES) is considered a key social determinant of depression (Maselko, 2017). In LMIC, household assets, educational attainment, and financial strain are commonly used to capture SES as these are considered most relevant to the processes of social stratification (Howe et al., 2012). Using these indicators, epidemiological studies in LMIC have shown that lower SES is associated with increased rates of depression (Lund and Cois, 2018; Maselko et al., 2017). However, although the greatest health disparities are observed in LMIC, little research has been conducted in these setting, instead it is focused on high-income countries (Polderman et al., 2015; Saxena et al., 2006). Reducing this burden requires a better understanding of the causal relationships underlying the observed association between socioeconomically disadvantaged circumstances and depression.

Two principal pathways are thought to underlie the observed associations between SES and mental health outcomes; social causation and social selection (Dohrenwend et al., 1992). The social causation hypothesis posits that exposure to the adverse social and economic conditions associated with lower SES (such as poor environmental conditions, material and social deprivation, and increased exposure to adverse and stressful life events) increases the risk for mental health conditions. The social selection hypothesis suggests that individuals with mental health disorders are more likely to drift into or remain in lower SES levels due to disability, reduced economic productivity, loss of employment, increased health expenditure, and stigma as a result of their illness. A recent longitudinal study in a nationally representative sample from South Africa found evidence for a reciprocal relationship between SES and depression, suggesting social causation and social selection act simultaneously to reinforce cycles of socioeconomic disadvantage and depression (Lund and Cois, 2018).

Associations between socioeconomic factors and depression could also arise because they share common causes. Genetic influences have been shown to be associated with

environmental exposures (Kendler and Baker, 2007). This is known as gene-environment correlation (rGE) and describes genetically influenced behaviour which can influence individuals' exposure to certain environments. Genetically informative research has provided evidence for significant genetic influence on depression (Sullivan et al., 2000) and indicators of SES (Ball et al., 2010; Hill et al., 2019; Rimfeld et al., 2018). If there is an overlap in the genetic factors associated with these traits, then part of the link between them could be explained by common genetic influences. Research has provided evidence for genetic overlap between socioeconomic indicators and depression (Hill et al., 2019), suggesting that part of the link between them could be explained by common genetic influences. These findings could be taken as support for the social selection hypothesis, in which affected individuals may be more likely to drift into or remain at lower SES levels, at least in part, based on genetically influenced traits and behaviours related to depression.

Socioeconomic conditions may also affect the relative importance of genetic and environmental influences on depression within a population. This is known as gene-environment interaction (GxE) and reflects a form of social causation whereby aetiological influences on a trait are moderated by context (Rutter et al., 2006). One study investigating GxE in depression in twins in the United States found non-shared environmental influences on internalising symptoms to be greater at lower levels of income (South and Krueger, 2011). This suggests that in environments with greater adversity, genetic effects on depression may be masked. This would lead to genetic effects being more clearly detected in enriched environments. In the context of HIC and LMIC, this would suggest that genetic effects would be easier to detect in HIC compared to LMIC. However, in another study, based in the United States, higher levels neighbourhood socioeconomic disadvantage were associated with greater genetic influences (Strachan et al., 2017). To date, genetically informative research has been largely restricted to high-income populations (Polderman et al., 2015). Bias towards Western populations is problematic because estimates of genetic and environmental sources of individual differences are specific to a population at a particular time. Studies conducted in different countries show modest evidence of differences in aetiology of depression, however, to date there has been no study which directly addresses cross-country variability by comparing different heritability estimates across multiple cohorts in different countries (Ball et al., 2009; Hur, 2008; Sullivan et al., 2000; Zavos et al., 2020).

Given the increased levels of socioeconomic disadvantage and disease burden in LMIC, further exploration is needed to understand individual differences in socioeconomic variables and their influence on depression. Here, we focus on Sri Lanka due to the under-representation of the Global South in genetically informed research and the opportunity presented by an existing twin study in the Colombo District. Sri Lanka serves a distinctive context for population mental health research due to its exposure to a major tsunami in 2004 and a prolonged civil war from 1983 to 2009. These events, tied to ethnic divisions and economic instability, were pervasive stressors impacting mental health at the population level. Examining the influence of socioeconomic factors on depression in Sri Lanka allows for the development of interventions mindful of the local and/or regional context.

In a population-based sample of Sri Lankan twins and singletons, we investigated (1) associations between socioeconomic indicators and depression symptoms; (2) the role of gene-environment correlation (rGE) in indicators of SES, if significant genetic influences on SES are observed, then this could reflect a form of social selection; and (3) whether socioeconomic indicators moderate the genetic and environmental influences on depression symptoms (GxE). If a significant interaction between socio-economic indicators and genetic influences on depression are observed, then this could reflect a form of social causation whereby certain socioeconomic contexts moderate aetiological influences on susceptibility to depression.

3.3. Methods

3.3.1. Sample

The Colombo Twin and Singleton Study (CoTASS) is a population-based study that took place between 2005-2007 in Colombo, Sri Lanka, including 4,009 twins (of which 1,954 were identified as complete twin pairs) and 2,019 singletons (Siribaddana et al., 2008). The initial participation rate was 91% among eligible twins and 87% among singletons. This study uses data from COTASS-2, a follow-up study conducted between 2012-2015. In COTASS-2, questionnaire data was available from 3934 twins (*N* = 2899) and singletons (*N* = 1035), comprising 76.4% of the original COTASS-1 sample (Jayaweera et al., 2018). The sample were 57.6% female, and the mean age was 42.8 years. Written informed consent was obtained from

all participants. Participants were offered 750 LKR (approximately £3.50) upon completion of one or more study components to compensate for time and inconvenience. Full details of the COTASS-2 study are described in Jayaweera et al. (2018). The study received ethical approval from the Faculty of Medical Sciences University of Sri Jayewardenepura Ethical Review Committee (USJP ERC; reference number: 596/11) and from the Psychiatric, Nursing and Midwifery Research Ethics Subcommittee, King's College London, UK (reference number: PNM/10/11-124).

3.3.2. Interview measures

Questionnaire data were collected by trained field research assistants. Interviews lasted 1-2 hours and were typically conducted in participants' homes. Questionnaires were translated into Sinhalese by a panel of health professionals fluent in both Sinhala and English. Translations were cross-culturally adapted in wording to best describe questionnaires in their meaning (Sumathipala and Murray, 2000).

3.3.2.1. Socioeconomic status (SES) indicators

Standard of living. Questionnaire items relating to housing conditions, ownership of household appliances and access to transport were used to index standard of living (see Supplementary Table S1). Composite standard of living scores were created by taking the sum of the items. Scores ranged between 1 to 17, with higher values indexing higher standard of living.

Educational attainment. Participants were asked to report their level of educational attainment. Response values ranged from 0 (*no education*) to 6 (*university or higher*).

Financial strain. To measure financial strain, participants were asked "how well do you feel you are managing financially these days?". Response options were based on a five-point scale ranging from "finding it very difficult to make ends meet" to "living comfortably".

3.3.2.2. Depression symptoms

The Revised Beck Depression Inventory (BDI-II) was used to measure depression symptoms and severity in the past two weeks (Beck et al., 1996). The BDI-II is a self-report questionnaire consisting of 21 items. For each item, four response options arranged in increasing severity are presented on a 4-point scale (0-3). Item-level scores were summed to create a composite score. Higher total scores indicated greater severity of depression symptoms. The BDI-II is a reliable measure of depression and has been previously validated in the Sri Lankan population (Rodrigo et al., 2015).

3.3.2.3. Zygosity

Zygosity was ascertained in CoTASS-1 using a self-report questionnaire measure of similarity (Siribaddana et al., 2008). If zygosity was missing in CoTASS-1, it was replaced with zygosity information collected using the same measure in CoTASS-2 (*n* = 88). Zygosity characteristics are in line with the usual distribution seen in population studies, with slightly more MZ versus DZ twin pairs, and opposite sex pairs being the largest group (Jayaweera et al., 2018).

3.3.3. Statistical analysis

We conducted a series of analyses to obtain the following estimates: (1) associations between SES indicators and depression symptoms; (2) estimates from univariate twin analyses; (3) estimates from biometric bivariate moderation (GxE) analyses. All analyses were conducted in R v.4.0.2 (https://www.R-project.org/; R Core Team, 2020).

3.3.3.1. Phenotypic associations

Linear regression analyses were performed to assess phenotypic associations between sociodemographic variables, SES indicators, and depression symptoms. Analyses were clustered using the 'lm.cluster' function in the 'miceadds' package (https://CRAN.R-project.org/package=miceadds) (Robitzsh and Grund, 2021) which returns clustered standard errors to account for the non-independence of twins in the sample.

3.3.3.2. Twin model fitting

Twin design

The twin design compares intra-class correlations of identical (monozygotic, MZ) and non-identical (dizygotic, DZ) twin pairs to estimate the contribution of genetic and environmental factors to observed phenotypic variance in a trait and/or covariance between traits (Rijsdijk and Sham, 2002). The classical twin method is based on the following assumptions: (1) MZ twins share 100% of their genes and DZ twins share on average 50% of their segregating genes (i.e., genes that differ between individuals); (2) MZ and DZ twin pairs share environmental influences common to both twins in the same family to the same extent ('shared environment'); and (3) MZ and DZ twin pairs differ from one another due to exposure to environmental factors which are unique to the individual ('non-shared environment'). The twin model attributes the similarity of reared-together twins to additive genetic (A) factors and shared environmental (C) factors that are common to both twins in the same family. The correlation between twins' shared environment is assumed to be 1 for both MZ and DZ pairs. The differences between MZ and DZ twin pairs is attributed to non-shared environmental influences (E) which are unique to the individual. By comparing differences in correlations between MZ and DZ twin pairs and linking these back to the model of the expected correlations (A+C for MZ pairs and 0.5A+C for DZ pairs), it is possible to establish the role of genetic and environmental influences. If MZ twins are more correlated on a trait than DZ twins, then genetic influences are assumed. Shared environmental influences are assumed if the DZ twin correlation is greater than half of MZ twin pairs. The extent to which MZ twins differ on a trait indicates non-shared environmental influences and measurement error.

Univariate ACE models

Structural equation model-fitting analyses were performed to estimate the relative contribution of additive genetic (A), shared environment (C), and non-shared environment (E) factors to the variation in SES indicators and depression symptoms. First, a heterogeneity ACE model was fit to the data in which the A, C and E parameters are estimated separately for

males and females allowing for quantitative sex differences. To test for variance differences between males and females, a scalar effects model was then performed, which allows only phenotypic variance differences between males and females but equates A, C and E in males and females. Last, a homogeneity model was fitted, in which scalar effects were dropped and all parameters were held equal for males and females. The relative fit of models allowing for different types of sex differences (i.e., quantitative and variance differences between the sexes) and no sex differences were compared to assess which model best describes the data.

Biometric moderation (GxE) models

Bivariate biometric moderation models (Purcell, 2002) were used to investigate whether SES indicators moderate the aetiology of depression. Modelling of biometric moderation in a structural equation framework allows for different ACE estimates for subgroups in the population with a certain standing on a moderator variable (Purcell, 2002; van der Sluis et al., 2012) (Figure 1). This model is an extension of a bivariate decomposition in which the variance in two variables, and the covariance between them, is partitioned into genetic, shared environmental and non-shared environmental effects. In the moderation model, the moderation effects are modelled directly on the path loadings of the ACE variance components unique to the trait, as well as the variance components shared between the trait and moderator (Figure 1). As such, it is possible to simultaneously model: (1) shared genetic and environmental effects between the moderator and trait, (2) the moderation of the genetic and environmental variance components shared between the moderator and trait, and (3) the moderation of the variance components unique to the trait. To test for the significance of moderation, we compared each full moderation model with a no-moderation model, in which all moderation parameters were dropped (b_{ac}, $b_{cc,}$ $b_{ec,}$ b_{au}, b_{cu} and b_{eu} constrained to zero). We then tested whether moderation on the individual ACE variance components was significant by examining the 95% confidence intervals (CI) around the moderated parameters from the full model. We did not fit constrained sub models to test for the significance of individual moderated parameters because omission of moderation effects by fixing them to 0 can bias estimation of parameters (e.g., dropping ßc can inflate ßa, and there are issues of specificity in distinguishing between ßa and ßc) (Van Hulle and Rathouz, 2015).

All SEM analyses were conducted using the open-source package OpenMx (Neale et al., 2016). OpenMx uses full-information maximum-likelihood to estimate model parameters. Study variables were age (in years) and sex corrected prior to model fitting and standardised residuals were used. The residual score for depression symptoms was log transformed to reduce positive skew. Models were fitted using full-information maximum-likelihood estimation and compared using likelihood ratio testing (differences in -2 Log-likelihood and associated degrees of freedom, which is x^2 distributed) and the Akaike's Information Criterion (AIC) (Akaike, 1987).

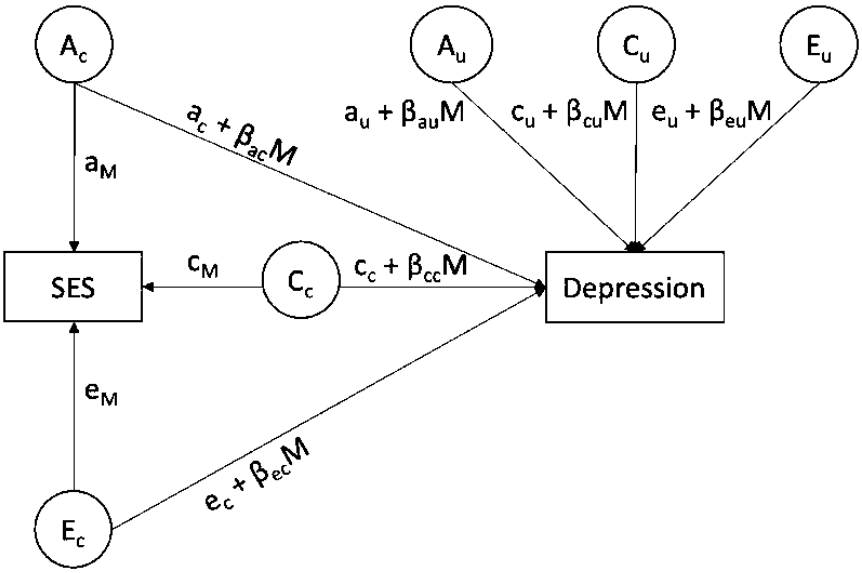


Figure 1. Bivariate moderation model shown for only one member of a twin pair as proposed by Purcell (2002). Ac, Cc and Ec are the variance components common to the moderator and the trait. Au, Cu, and Eu are the variance components unique to the trait. Path loadings for the moderator are denoted by a_M, c_M, and e_M. The cross-paths connecting the moderator to the trait consist of loadings that are unrelated to the moderator: a_c, c_c, and e_c, and cross-loadings that depend on the moderator via weights: b_{ac}, b_{cc}, and b_{ec}. The path loadings unique to the trait consist of elements unrelated to the moderator a_u, c_u, and e_u, and elements that depend on the moderator via weights b_{au}, b_{cu}, and b_{eu}. b coefficients index the direction and magnitude of moderation. The total variance of the trait can be calculated as follows: $Var(T|M) = (a_c + b_{ac}M)^2 + (a_u + b_{au}M)^2 + (c_c + b_{cc}M)^2 + (c_u + b_{cu}M)^2 + (e_c + b_{ec}M)^2 + (e_u + b_{eu}M)^2$.

3.4. Results

3.4.1. Associations between SES indicators, sociodemographic characteristics and depression symptoms

Adjusted associations between sociodemographic variables and SES indicators are shown in Table 1 (see Supplementary Table S2 for unadjusted b coefficients). Females reported lower levels of standard of living and financial strain compared to males. Older age and being from a minority ethnic group (Tamil or Muslim) was associated with lower educational attainment. Those who had been previously married reported lower educational attainment compared to married individuals. Living in non-urban areas was associated with lower standard of living and higher financial strain. The indicators of SES were associated with one another

(Supplementary Table S2), and associations remained significant after adjustment for other sociodemographic factors (Table 1).

The mean score of BDI-II depression symptoms reported in the current sample was 4.86. Higher depression scores were observed in females compared to males (b = 1.48, 95% CI [1.08, 1.88]). Higher depression scores were significantly associated with lower standard of living, lower educational attainment, and financial strain (Table S3). These associations remained significant after adjusting for sex and age (Table S3).

3.4.2. Univariate ACE model fitting

Twin correlations and ACE parameter estimates from the best-fitting models are presented in Table 2. Fit statistics from the univariate ACE model-fitting analyses are shown in Supplementary Table S4. Significant genetic influences were observed for standard of living in females, and for educational attainment in both males and females. Moderate shared and non-shared environmental influences were also apparent for both standard of living and educational attainment. Twin correlations for financial strain were higher in male DZs compared to MZs and the univariate ACE model did not fit the data, suggesting that this variable will not conform to any genetic model. We therefore did not pursue further biometric model fitting with this variable. Variance in depression was explained by genetic and non-shared environmental influences.

3.4.3. Biometric moderation (GxE) model fitting

Bivariate biometric moderation models were applied to test whether the aetiology of depression symptoms was moderated by (1) standard of living or (2) educational attainment. Scalar sex differences for depression symptoms were modelled. We did not examine whether standard of living moderated the aetiology of depression symptoms separately for males and females due to limited sample size. We observed significant moderation as a function of each SES indicator (Table S5), however, could not determine whether this was due to genetic or environmental moderation. Moderated parameter estimates derived from each of the full moderation models are shown in Table 3.

Table 1. Associations between socioeconomic status indicators and sociodemographic characteristics.

		Standard of living	Educational attainment	Financial strain
	N	Adjusted b	Adjusted b	Adjusted b
Sex				
Male (ref)	1681 (42.4%)			
Female	2288 (57.6%)	-0.10 (-0.15, -0.04)**	0.11 (0.06, 0.17)**	-0.07 (-0.13, -0.01)*
Age				
19-29 (ref)	853 (21.5%)			
30-39	1012 (25.5%)	0.00 (-0.09, -0.10)	-0.26 (-0.35, -0.17)**	-0.02 (-0.11, 0.08)
40-49	825 (20.8%)	0.15 (0.05, 0.25)**	-0.42 (-0.51, -0.32)**	-0.14 (-0.25, -0.03)*
50-59	665 (16.8%)	0.23 (0.12, 0.34)**	-0.53 (-0.63, -0.43)**	-0.07 (-0.18, 0.04)
60-69	376 (9.5%)	0.22 (0.10, 0.35)**	-0.59 (-0.71, -0.48)**	-0.12 (-0.26, 0.01)
>70	203 (5.1%)	0.22 (0.04, 0.40)*	-0.63 (-0.79, -0.47)**	0.13 (-0.04, 0.31)
Ethnicity				
Sinhala (ref)	3647 (91.9%)			
Tamil	120 (3.0%)	-0.22 (-0.44, -0.00)	-0.24 (-0.44, -0.04)*	-0.02 (-0.23, 0.19)
Muslim	150 (3.8%)	0.26 (0.11, 0.40)**	-0.49 (-0.61, -0.36)**	0.01 (-0.13, 0.15)
Other Minority	16 (0.4%)	0.38 (0.15, 0.61)**	-0.16 (-0.56, 0.24)	-0.41 (-0.91, 0.09)
Marital Status				
Married (ref)	2838 (71.5%)			
Previously Married	329 (8.3%)	-0.05 (-0.17, 0.06)	-0.21 (-0.32, -0.11)**	-0.09 (-0.22, 0.05)
Never Married	763 (19.2%)	-0.01 (-0.10, 0.09)	0.17 (0.08, 0.25)**	0.01 (-0.09, 0.10)
Urbanicity				

Urban (ref)	2390 (60.2%)			
Rural	532 (13.4%)	-0.29 (-0.38, -0.21)**	-0.02 (-0.12, 0.07)	0.31 (0.23, 0.39)**
Mixed	826 (20.8%)	-0.15 (-0.22, -0.08)**	-0.11 (-0.18, -0.04)**	0.22 (0.14, 0.29)**
Outside Colombo	221 (5.6%)	-0.21 (-0.33, -0.10)**	0.17 (0.05, 0.29)**	-0.03 (-0.17, 0.11)
Standard of Living				
Mean (SD)	14.1 (2.64)	-	0.15 (0.14, 0.16)**	0.12 (0.11, 0.14)**
Educational attainment				
No education (ref)	47 (1.2%)		-	
Grade 1-5	274 (6.9%)	0.10 (-0.31, 0.51)	-	-0.22 (-0.59, 0.16)
Grade 6 0/Ls	1757 (44.3%)	0.47 (0.07, 0.86)*	-	0.02 (-0.33, 0.36)
Passed O/Ls	632 (15.9%)	0.91 (0.51, 1.31)**	-	0.25 (-0.10, 0.61)
Up to/ passed A/Ls	929 (23.4%)	1.19 (0.79, 1.59)**	-	0.29 (-0.06, 0.64)
University /higher	276 (7.0%)	1.51 (1.10, 1.91)**	-	0.30 (-0.07, 0.66)
Financial Strain				-
Very difficult to make ends meet (ref)	121 (3.0%)			-
Difficult to make ends meet	284 (7.2%)	0.50 (0.25, 0.76)**	0.01 (-0.15, 0.18)	-
Just about getting by	547 (13.8%)	0.62 (0.37, 0.87)**	0.05 (-0.11, 0.21)	-
Doing alright	2616 (65.9%)	1.01 (0.77, 1.24)**	0.24 (0.08, 0.39)	-
Living comfortably	365 (9.2%)	1.45 (1.20, 1.70)**	0.54 (0.36, 0.72)**	-

Note. Linear regressions were conducted using standardised outcome variables and clustered standard errors to account for non-independence of twins in the sample. Adjusted b coefficients were calculated after including all other socio-demographic variables in the table.

*p < 0.05; **p < 0.01

Table 2. Twin correlations and univariate ACE estimates for standard of living, educational attainment, financial strain, and depression.

	Standard of living	Educational attainment	Financial strain	Depression
MZM	0.65 (0.58, 0.71)	0.68 (0.61, 0.73)	0.38 (0.25, 0.48)	0.28 (0.13, 0.40)
DZM	0.67 (0.57, 0.74)	0.58 (0.47, 0.67)	0.69 (0.59, 0.75)	0.23 (0.07, 0.37)
MZF	0.63 (0.57, 0.68)	0.74 (0.70, 0.78)	0.50 (0.42, 0.57)	0.36 (0.26, 0.46)
DZF	0.56 (0.46, 0.63)	0.54 (0.45, 0.62)	0.40 (0.29, 0.50)	0.21 (0.07, 0.34)
DZOS	0.50 (0.41, 0.57)	0.47 (0.39, 0.54)	0.39 (0.28, 0.48)	0.12 (0.01, 0.22)

	A	C	E
Standard of living			
Male	0.03 (0.00, 0.10)	0.63 (0.55, 0.70)	0.34 (0.29, 0.40)
Female	0.18 (0.03, 0.36)	0.46 (0.29, 0.59)	0.37 (0.32, 0.43)
Educational attainment	0.40 (0.28, 0.52)	0.32 (0.20, 0.42)	0.28 (0.25, 0.32)
Depression	0.32 (0.13, 0.40)	0.01 (0.00, 0.18)	0.67 (0.60, 0.76)

Note. MZM = monozygotic male, DZM = dizygotic male, MZF = monozygotic female, DZF = dizygotic female, DZOS = dizygotic opposite sex. A = additive genetic, C = shared environmental, and E = non-shared environmental influences.

Table 3. Genetic and environmental parameter estimates for depression moderated by standard of living and educational attainment.

	Standard of living	Educational attainment
Parameter		
bau	0.01 (-0.05, 0.06)	-0.04 (-0.09, 0.01)
bac	-0.01 (-0.10, 0.07)	-0.02 (-0.08, 0.05)
bcu	-0.15 (-0.23, 0.23)	0.09 (-0.20, 0.20)
bcc	0.00 (-0.06, 0.06)	-0.02 (-0.09, 0.05)
beu	-0.04 (-0.08, 0.00)	-0.02 (-0.06, 0.01)
bec	0.02 (-0.02, 0.06)	0.04 (0.00, 0.09)

Note. Parameter estimates are derived from the full moderation (GxE) model. bau, bcu, and beu are the moderated genetic, shared environmental, and non-shared environmental path coefficients unique to depression symptoms. bac, bcc, and bec are the moderated genetic, shared environmental, and non-shared environmental path coefficients common to the moderator (i.e., standard of living or educational attainment) and depression symptoms.

3.4.4. Depression symptoms moderated by standard of living

Figure 2a shows the unstandardised variance in depression symptoms moderated by standard of living. Lower standard of living was associated with greater variance in depression symptoms compared to higher levels of standard of living. Dropping moderation parameters

resulted in a significant decrease in fit compared to the full moderation model (Table S5). Figure 2a suggests that genetic influences unique to depression symptoms increased with higher standard of living. Shared and non-shared environmental influences unique to depression were greater at lower standard of living. However, no single variance component appeared significantly moderated by standard of living when confidence intervals around the moderation terms were inspected (Table 3).

3.4.5. Depression symptoms moderated by educational attainment

Figure 2b shows the unstandardised variance in depression symptoms moderated by educational attainment. Total variance in depression symptoms was greater at lower levels of educational attainment. Dropping all moderation parameters resulted in a significant decrease in fit compared to the full moderation model (Table S5). Figure 2b suggests that genetic influences unique to depression symptoms were lower at high, compared to low levels of educational attainment. Environmental influences unique to depression symptoms do not appear to vary greatly as a function of educational attainment. However, this should be considered indicative as confidence intervals indicated that no individual variance component was significantly moderated by educational attainment (Table 3).

3.4.6. Post-hoc analyses

Post-hoc model fitting analyses were performed to assess the moderating effects of SES indicators on (1) the ACE variance components shared between depression and each SES indicator (b_{ac}, b_{cc} and b_{ec} constrained to zero), and (2) the variance components unique to depression symptoms (b_{au}, b_{cu} and b_{eu} constrained to zero). Results showed that the moderated ACE variance components unique to depression symptoms could not be dropped without a significant worsening of fit to the data compared with the full moderation models (Table S5). This suggests presence of moderation on the variance components unique to depression and is consistent with the moderation effects observed in Figure 2. No individual variance component appeared significantly moderated by standard of living or educational attainment when confidence intervals around the moderation terms were inspected. Further, post-hoc *phenotypic* moderation model fitting analyses were performed to assess the moderating effects of SES indicators on (1) the phenotypic variance shared between depression and each SES indicator (Bc constrained to zero), and (2) the phenotypic variance

unique to depression symptoms (Bu constrained to zero). This allowed us to estimate one overall Beta-c (with 95% CI) and one overall Beta-u (with 95% CI).

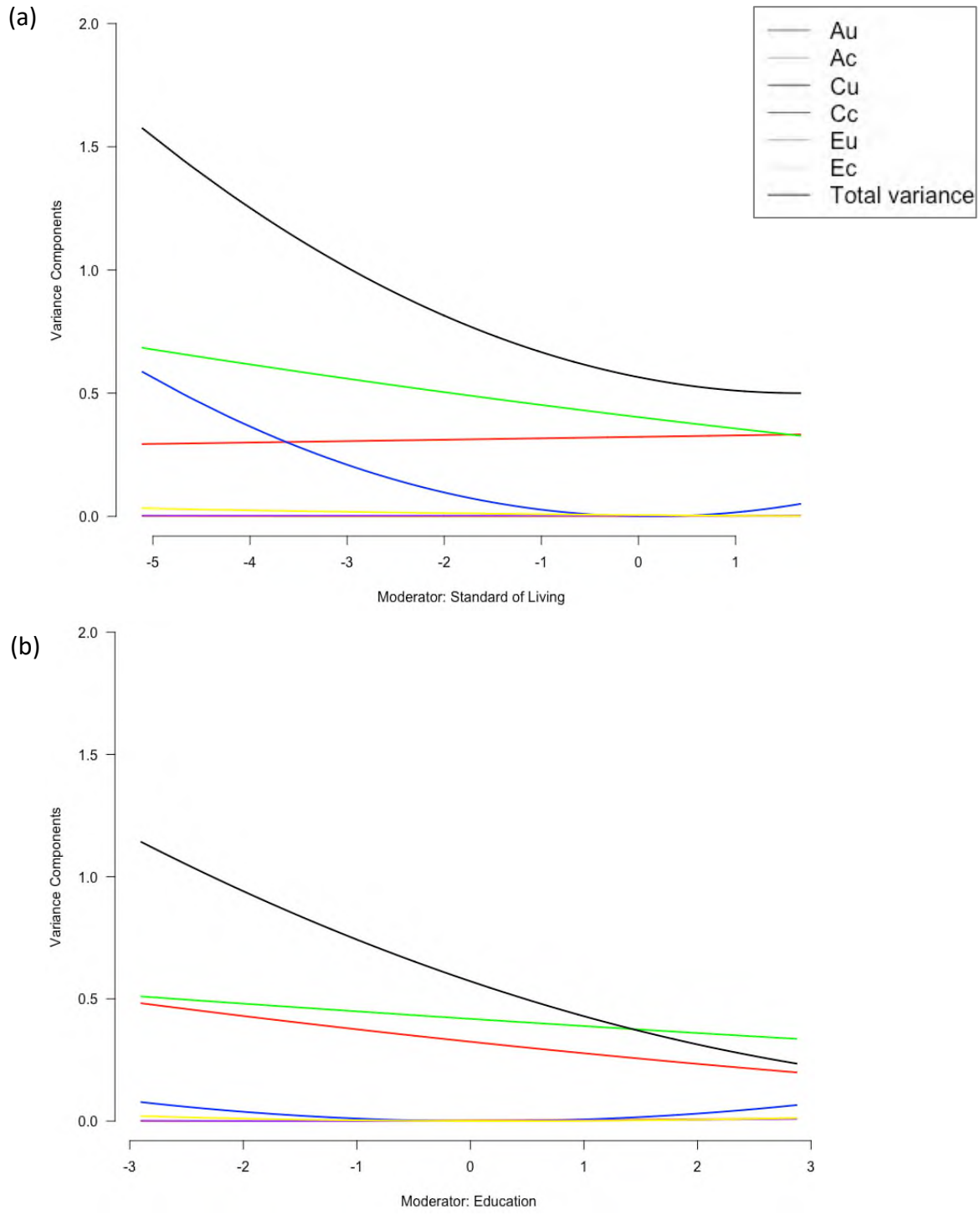


Figure 2. Variance in depression symptoms moderated by standard of living (a) and educational attainment (b). Au, Cu, and Eu are the genetic, shared environmental, and non-shared environmental variance components unique to depression. Ac, Cc and Ec are the are the genetic, shared environmental, and non-shared environmental variance components common to the socioeconomic moderator and depression.

Results showed that the moderated variance unique to depression symptoms could not be dropped without a significant worsening of fit to the data compared with the full moderation models (Table S6). Examination of the confidence intervals around the moderation parameter estimates from the full model showed significant moderation of the variance unique to depression symptoms (Bu = -0.04, 95% CI [-0.06, -0.02]). This suggests presence of moderation on the overall variance unique to depression and is consistent with the moderation effects observed in Figure 2. Post-hoc bivariate analyses were also conducted to assess genetic and environmental correlations between each SES indicator and depression symptoms. Significant genetic correlations between depression symptoms and both standard of living and educational attainment were observed (Table S7).

3.5. Discussion

The current study sought to examine the relationship between socioeconomic factors and depression symptoms using genetically informative data from a population-based sample of twins and singletons in Sri Lanka. In line with previous findings (Maselko et al., 2017), we found moderate phenotypic associations between the different socioeconomic indicators. Lower standard of living, poor educational attainment and financial strain were independently associated with higher depression symptoms, consistent with previous research indicating that individuals with lower SES are at increased risk for depression (Lund and Cois, 2018; Maselko, 2017; Maselko et al., 2017). Results provide support for both social selection (significant genetic influences on SES indicators [rGE]) and preliminary support for social causation (GxE).

Sex differences were identified in the aetiology of standard of living, with evidence of a small contribution of genetic influences on standard of living in females but not in males. Genetic influences on standard of living in females may be accounted for lower variation in environmental exposures due to cultural gender limitations. For example, the majority of working age females in Sri Lanka are not in salary-based employment (~73%) and economically inactive (Department of Census and Statistics Sri Lanka, 2019). The main reason reported is caregiver, family work and housework activities (Department of Census and Statistics Sri Lanka, 2019). Environmental influences explained the majority of variance for standard of living. Our results are broadly in line with previous research using an earlier wave

of the COTASS sample (Ball et al., 2010). Our finding that 44-63% of the variance in standard of living was due to environmental factors shared within the family (C) contrast with reports of relatively small or zero shared environmental effects on socioeconomic indicators in studies from adults in HIC (Rimfeld et al., 2018). This could be explained by differences in socio-cultural norms such as higher prevalence of extended, multigenerational, family households and greater importance given to family-based networks in LMIC (Maselko, 2017). Larger environmental variation in standard of living could also indicate less equal access or opportunity in employment sectors in Sri Lanka compared to in countries where higher heritability estimates for socioeconomic factors have been reported (Rimfeld et al., 2018). Informal employment is estimated to account for 66.7% of total employment in Sri Lanka compared to 18% in HIC (Bonnet et al., 2019; Department of Census and Statistics Sri Lanka, 2019). Those who are employed in the informal economy face multiple challenges, such as low job security and difficult working conditions, and informal work is often undertaken due to absence of other means of livelihood (Bonnet et al., 2019). Given that traits and behaviours related to SES are substantially genetic in origin, greater equality of opportunity means that environmental inequalities, such as privilege or prejudice, have less impact on outcomes. Individual differences in socioeconomic factors that remain after systemic environmental inequalities are reduced are to a greater extent due to genetic differences.

Educational attainment was moderately heritable with a significant contribution of shared and non-shared environmental influences in males and females. Heritability of educational attainment in both sexes may be indicative of gene-environment correlation. Higher heritability for educational attainment could reflect that the education system is more meritocratic in Sri Lanka than other aspects of the socioeconomic context (Rimfeld et al., 2018). Primary and secondary education in Sri Lanka is free and enrolment in secondary education is 91% for both genders (UNESCO Institute for Statistics, 2018). This is consistent with research demonstrating higher heritability for educational attainment in societies with greater equality in educational opportunities (Rimfeld et al., 2018).

We found evidence that SES indicators moderated the aetiology of depression. However, we did not find significant moderation of the individual ACE variance components unique to depression, contrary to prior work (South and Krueger, 2011; Strachan et al., 2017). Our most

consistent finding was that total variance in depression symptoms was greater among lower-SES individuals, which was driven by greater genetic and environmental variance components unique to depression at lower levels of SES, without the ability to detect moderation of each component individually as significant. This is partially consistent with findings from previous GxE studies (South and Krueger, 2011; Strachan et al., 2017) and provides evidence that is consistent with the notion that social causation processes play a role in the observed association between depression and SES across different populations (Dohrenwend et al., 1992; Lund and Cois, 2018; South and Krueger, 2011). Our results also showed that the shared variance components between socioeconomic indicators and depression symptoms were zero across the entire SES distribution. This suggests that socioeconomic factors may have a main effect on the aetiology of depression independently of shared aetiological influences.

Strengths and limitations

A strength of our study is the use of a large representative population-based twin and singleton sample based in Sri Lanka, especially given the limited availability of genetically informative data in LMIC populations (Polderman et al., 2015). We used different socioeconomic indicators intended to capture different aspects of the socioeconomic context. Thus, we were able to examine and compare individual differences in socioeconomic outcomes and their role in the aetiology of depression symptoms. In addition, the wide availability of asset index data and educational attainment in many studies and comparable data across multiple countries is an important strength because it facilitates comparative research. Some limitations should be considered. First, self-reported socioeconomic factors and depression symptoms could be underreported due to the sensitive and/or private nature and stigma associated with reporting them (Lorant et al., 2007). Second, we did not investigate whether the pattern of moderating effects by socioeconomic conditions vary over age. Differences in moderation could be expected given that the aetiology of mental health problems changes across development (Hannigan et al., 2017). Future studies could seek to explore how the dynamics of different socioeconomic conditions relate to aetiologic moderation on depression symptoms across age. Third, while the sample is representative of individuals living in Colombo District, it may not be representative of different regions of Sri Lanka. Additionally, the small sample sizes for ethnic minorities in this study, such as Sri Lankan

Tamils (who live predominantly in the north and east of the island and are disproportionately impacted by poverty resulting from ethnic tensions and the lasting economic effects of the civil war), prevent a detailed investigation into how socioeconomic status might uniquely influence depression within specific ethnic groups. This limits the generalisability of our findings to diverse socio-cultural contexts across Sri Lanka. Lastly, the twin method rests on certain assumptions that when unmet may challenge the validity of the results (Rijsdijk and Sham, 2002).

3.5.1. Conclusion

The present study extends our understanding of the relationship between socioeconomic factors and depression symptoms using data from a representative twin and singleton population study based in Colombo, Sri Lanka. Shared and non-shared environmental influences accounted for the majority of variance in standard of living, whereas educational attainment showed moderate heritability. Socioeconomic indicators moderated the variance unique to depression symptoms, consistent with previous investigations in samples drawn from different social, economic, and cultural contexts. However, we were unable to determine whether this was due to genetic or environmental moderation. This is the first study to use bivariate moderation modelling to investigate whether socioeconomic factors moderate aetiological influences on depression symptoms in a South Asian population. This study has implications for future cross-cultural investigations of the mechanisms underlying associations between socioeconomic factors and depression symptoms and has the potential to inform intervention strategies to reduce social disparities in depression.

3.6. Acknowledgements

This research was funded in whole, or in part, by the Wellcome Trust [Grant number 093206/Z/10/Z]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manscript version arising from this submission. TAM is supported by a Wellcome Trust Senior Research Fellowship (220382/Z/20/Z. IB is supported by the UK Economic and Social Research Council (ESRC) and King's College London member of the ESRC Doctoral Training Partnership in Interdisciplinary Social Science (ST11872). We gratefully acknowledge the contribution of the teams who have contributed immensely to the collection and preparation of the data. The data collection team includes: Maneesha

Jayaweera, Geetha Weerasekara, Ajith Weerasekara, Chandrika Wickramasuriya, Ravindra Perera, Aruna Walisundara, Nadun Perera, Thirumalayandi Jayakumar, Kasun Dias, Amila Sandaruwan, Nipuna Siribaddana, and Vidath Samarakkody. The IRD office and data entry team: Chamali Jayasinghe, Mekala Narangoda, Janani Marasinghe, Dr. Jeewani Hettiarachchi, Udeni Samanmali, Nilanthi Priyadharshani, Hansika Gamage, and Pavithra Lakmali. The IRD office and data entry team: Lasith Dissanayake, Lakshan Warnakula, Shavindya Abeygunaratne, Rosanthy Skandhakumar, Ravindi Madanayake, Thathmi Wathsara and Charuni Rajapaksha. Thank you to all the participants of the CoTASS-2 study.

3.7. Conflict of Interest

Matthew Hotopf declares research funding from Innovative Medicines Initiative (European Commission), Janssen, Lundbeck, MSD, Biogen and UCB. All other authors declare no competing interests.

3.8. References

Akaike, H., 1987. Factor analysis and AIC. Psychometrika 52, 317–332. https://doi.org/10.1007/BF02294359

Ball, H.A., Siribaddana, S.H., Sumathipala, A., Kovas, Y., Glozier, N., McGuffin, P., Hotopf, M., 2010. Environmental exposures and their genetic or environmental contribution to depression and fatigue: a twin study in Sri Lanka. BMC Psychiatry 10, 13. https://doi.org/10.1186/1471-244X-10-13

Ball, H.A., Sumathipala, A., Siribaddana, S.H., Kovas, Y., Glozier, N., McGuffin, P., Hotopf, M., 2009. Genetic and environmental contributions to depression in Sri Lanka. Br J Psychiatry 195, 504–509. https://doi.org/10.1192/bjp.bp.109.063529

Beck, A.T., Steer, R.A., Brown, G., 1996. Beck Depression Inventory–II [WWW Document]. URL https://doi.org/10.1037/t00742-000 (accessed 5.13.21).

Bonnet, F., Vanek, J., Chen, M., 2019. Women and Men in the Informal Economy: A Statistical Brief. WIEGO, Manchester, UK.

Department of Census and Statistics Sri Lanka, 2019. Sri Lanka Labour Force Survey - Annual Report 2019 (Annual Report). Ministry of Finance.

Dohrenwend, B.P., Levav, I., Shrout, P.E., Schwartz, S., Naveh, G., Link, B.G., Skodol, A.E., Stueve, A., 1992. Socioeconomic status and psychiatric disorders: the causation-selection issue. Science 255, 946–952. https://doi.org/10.1126/science.1546291

Hannigan, L.J., Walaker, N., Waszczuk, M.A., McAdams, T.A., Eley, T.C., 2017. Aetiological Influences on Stability and Change in Emotional and Behavioural Problems across Development: A Systematic Review. Psychopathology Review a4, 52–108. https://doi.org/10.5127/pr.038315

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

Howe, L.D., Galobardes, B., Matijasevich, A., Gordon, D., Johnston, D., Onwujekwe, O., Patel, R., Webb, E.A., Lawlor, D.A., Hargreaves, J.R., 2012. Measuring socio-economic position for epidemiological studies in low- and middle-income countries: a methods of measurement in epidemiology paper. Int J Epidemiol 41, 871–886. https://doi.org/10.1093/ije/dys037

Hur, Y.-M., 2008. Sex Differences in Genetic and Environmental Contributions to Depression Symptoms in South Korean Adolescent and Young Adult Twins. Twin Research and Human Genetics 11, 306–313. https://doi.org/10.1375/twin.11.3.306

Jayaweera, K., Aschan, L., Pannala, G., Adikari, A., Glozier, N., Ismail, K., Pariante, C.M., Rijsdijk, F., Siribaddana, S., Zavos, H.M.S., Zunszain, P.A., Sumathipala, A., Hotopf, M., 2018. The Colombo Twin and Singleton Follow-up Study: a population based twin study of psychiatric disorders and metabolic syndrome in Sri Lanka. BMC Public Health 18, 145. https://doi.org/10.1186/s12889-017-4992-2

Kendler, K.S., Baker, J.H., 2007. Genetic influences on measures of the environment: a systematic review. Psychol Med 37, 615–626. https://doi.org/10.1017/S0033291706009524

Lorant, V., Demarest, S., Miermans, P.-J., Van Oyen, H., 2007. Survey error in measuring socio-economic risk factors of health status: a comparison of a survey and a census. International Journal of Epidemiology 36, 1292–1299. https://doi.org/10.1093/ije/dym191

Lund, C., Cois, A., 2018. Simultaneous social causation and social drift: Longitudinal analysis of depression and poverty in South Africa. Journal of Affective Disorders 229, 396–402. https://doi.org/10.1016/j.jad.2017.12.050

Maselko, J., 2017. Social Epidemiology and Global Mental Health: Expanding the Evidence from High-Income to Low- and Middle-Income Countries. Curr Epidemiol Rep 4, 166–173. https://doi.org/10.1007/s40471-017-0107-y

Maselko, J., Bates, L., Bhalotra, S., Gallis, J.A., O'Donnell, K., Sikander, S., Turner, E.L., 2017. Socioeconomic status indicators and common mental disorders: Evidence from a study of prenatal depression in Pakistan. SSM Popul Health 4, 1–9. https://doi.org/10.1016/j.ssmph.2017.10.004

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

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

Purcell, S., 2002. Variance Components Models for Gene–Environment Interaction in Twin Analysis. Twin Research and Human Genetics 5, 554–571. https://doi.org/10.1375/twin.5.6.554

Rijsdijk, F.V., Sham, P.C., 2002. Analytic approaches to twin data using structural equation models. Brief Bioinform 3, 119–133. https://doi.org/10.1093/bib/3.2.119

Rimfeld, K., Krapohl, E., Trzaskowski, M., Coleman, J.R.I., Selzam, S., Dale, P.S., Esko, T., Metspalu, A., Plomin, R., 2018. Genetic influence on social outcomes during and after the Soviet era in Estonia. Nat Hum Behav 2, 269–275. https://doi.org/10.1038/s41562-018-0332-5

Robitzsh, A., Grund, S., 2021. miceadds: Some Additional Multiple Imputation Functions, Especially for "mice."

Rodrigo, A., Kuruppuarachchi, K. a. L.A., Pathmeswaran, A., 2015. Validation of the Beck Depression Inventory II among the Sinhalese speaking population in Sri Lanka. Sri Lanka Journal of Psychiatry 6, 20–24. https://doi.org/10.4038/sljpsyc.v6i2.8076

Rutter, M., Moffitt, T.E., Caspi, A., 2006. Gene-environment interplay and psychopathology: multiple varieties but real effects. J Child Psychol & Psychiat 47, 226–261. https://doi.org/10.1111/j.1469-7610.2005.01557.x

Saxena, S., Paraje, G., Sharan, P., Karam, G., Sadana, R., 2006. The 10/90 divide in mental health research: trends over a 10-year period. Br J Psychiatry 188, 81–82. https://doi.org/10.1192/bjp.bp.105.011221

Siribaddana, S.H., Ball, H.A., Hewage, S.N., Glozier, N., Kovas, Y., Dayaratne, D., Sumathipala, A., McGuffin, P., Hotopf, M., 2008. Colombo Twin and Singleton Study (CoTASS): A description of a population based twin study of mental disorders in Sri Lanka. BMC Psychiatry 8, 49. https://doi.org/10.1186/1471-244X-8-49

South, S.C., Krueger, R.F., 2011. Genetic and environmental influences on internalizing psychopathology vary as a function of economic status. Psychol Med 41, 107–117. https://doi.org/10.1017/S0033291710000279

Strachan, E., Duncan, G., Horn, E., Turkheimer, E., 2017. Neighborhood deprivation and depression in adult twins: genetics and gene×environment interaction. Psychol Med 47, 627–638. https://doi.org/10.1017/S0033291716002622

Sullivan, P.F., Neale, M.C., Kendler, K.S., 2000. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry 157, 1552–1562. https://doi.org/10.1176/appi.ajp.157.10.1552

Sumathipala, A., Murray, J., 2000. New approach to translating instruments for cross-cultural research: a combined qualitative and quantitative approach for translation and consensus generation. International Journal of Methods in Psychiatric Research 9, 87–95. https://doi.org/10.1002/mpr.83

UNESCO Institute for Statistics, 2018. School enrollment, secondary (% net) - Sri Lanka | Data.

van der Sluis, S., Posthuma, D., Dolan, C.V., 2012. A Note on False Positives and Power in G × E Modelling of Twin Data. Behav Genet 42, 170–186. https://doi.org/10.1007/s10519-011-9480-3

Van Hulle, C.A., Rathouz, P.J., 2015. Operating Characteristics of Statistical Methods for Detecting Gene-by-Measured Environment Interaction in the Presence of Gene-Environment Correlation under Violations of Distributional Assumptions. Twin Res Hum Genet 18, 19–27. https://doi.org/10.1017/thg.2014.81

World Health Organization, 2017. Depression and Other Common Mental Disorders: Global Health Estimates.

Zavos, H.M.S., Dalton, B., Jayaweera, K., Harber-Aschan, L., Pannala, G., Adikari, A., Hatch, S.L., Siribaddana, S., Sumathipala, A., Hotopf, M., Rijsdijk, F.V., 2020. The relationship between independent and dependent life events and depression symptoms in Sri Lanka: a twin and singleton study. Soc Psychiatry Psychiatr Epidemiol 55, 237–249. https://doi.org/10.1007/s00127-019-01765-z

4. Evaluating bias associated with attrition in the Norwegian Mother, Father and Child Cohort Study

Supplementary materials for this chapter, as detailed in the text, are included in Appendix C.

4.1. Abstract

Background: Selection bias and attrition pose challenges to longitudinal cohort studies, potentially biasing estimates and impacting exposure-outcome associations. This study aimed to investigate the role of selective participation in the Mother, Father, and Child Birth Cohort (MoBa) and its impact on phenotypic and genetic associations. We examined factors associated with continued participation, assessed the bias introduced by selective participation on exposure-outcome associations, and explored the genetic underpinnings of continued participation and its relationship with maternal educational attainment and internalising symptoms.

Methods: We conducted regression analyses to identify factors associated with continued participation in MoBa and performed moderation analyses to evaluate the impact of selective participation on exposure-outcome associations. Biometric model-fitting analyses using an extended family design were employed to estimate the heritability of continued participation and to explore genetic correlations between educational attainment, internalising symptoms, and continued participation. To examine the impact of selection bias on genetic and environmental correlations, we assessed the aetiological overlap between maternal educational attainment and internalising symptoms in the full sample and in the subsample of individuals who returned the MoBa questionnaire at 8 years.

Results: Older age at birth, first-time mothers, and higher educational attainment were associated with increased rates of continued participation in MoBa, while later recruitment and higher levels of internalising symptoms were associated with attrition. Selective participation introduced bias in the association between educational attainment and internalising symptoms, suggesting the presence of collider bias. Genetic influences accounted for a significant proportion (34%) of the variance in continued participation, with non-shared environmental influences explaining the remaining variance. Genetic factors related to educational attainment or internalising symptoms were correlated with continued participation. The observed genetic correlation between educational attainment and internalising symptoms was similar in the full sample and the subsample of individuals who

returned the questionnaire at 8 years, suggesting minimal impact of selection bias on the observed genetic correlation.

Conclusion: Selective participation in MoBa was associated with socioeconomic and mental health factors, leading to bias in the phenotypic association between educational attainment and internalising symptoms. Genetic factors played a role in continued participation and showed correlations with educational attainment and internalising symptoms. However, our findings suggest that the observed genetic correlation between educational attainment and internalising symptoms is robust and not substantially influenced by selective attrition at follow-up assessments. Understanding the impact of selection bias and genetic influences on continued participation enhances our knowledge of cohort dynamics and informs strategies to mitigate bias in longitudinal studies.

4.2. Background

Selection bias is a pervasive problem for all scientific research that relies upon voluntary participation (Munafò et al., 2018; Nohr & Liew, 2018; Nunan et al., 2018). Bias through selective participation can arise when there are systematic differences between individuals who volunteer to participate in a research study and those who do not. Selection bias can occur at various stages in a study, including at recruitment, at follow-up or if a non-random subsample is selected for further investigation (e.g., based on exposure to a specific risk factor). Consequently, sample selection at baseline and subsequent follow-ups has the potential to bias estimates of disease prevalence and exposure-outcome associations. Sociodemographic, lifestyle and health factors have all been associated with the likelihood of individuals becoming or remaining participants in a study with those who are less advantaged and less healthy are often under-represented in studies. This has led to samples that are generally 'healthier and wealthier' than the intended study population (Dupuis et al., 2019; Goldberg et al., 2001; Graaf et al., 2000; Hara et al., 2002; Howe et al., 2013; Lamers et al., 2018; Nohr & Liew, 2018). When differences between participants and nonparticipants relate to the risk factors and outcomes under investigation, associations between risk and outcome may become biased. There are multiple ways in which selection bias can impact estimates, including collider bias, and omission of subpopulations. Resultant biases can induce associations where there is no causal effect, attenuate true causal effects, or reverse the sign of a causal effect (Munafò et al., 2018; Nohr & Liew, 2018; Smith, 2020).

Researchers have investigated selection bias in exposure-outcome association estimates either by comparing associations in the study sample with those in the target population or by comparing continuing participants with dropouts. The former approach requires data on relevant exposure and outcome characteristics for the entire target population (typically rare except, for example, through linkage to population and health registers). The latter requires only baseline data on all participants in a longitudinal study. As such, most studies of selection bias tend to examine biases attributable to differences between continuing participants and dropouts, rather than between participants and non-participants. That is, they focus on bias due to attrition rather than bias associated with initial participation/recruitment.

Of studies examining the impact of selection bias on exposure-outcome associations between health outcomes and established risk factors, most have reported that bias is minimal (R. M. Nilsen et al., 2009; Nohr et al., 2006; Nohr & Liew, 2018; Pizzi et al., 2011; Winding et al., 2014; Wolke et al., 2009). For example, research using data from the Medical Birth Registry of Norway found similar associations between study participants of the Norwegian Mother, Father, and Child Birth Cohort and the total population of women giving birth in Norway for several exposures and adverse pregnancy outcomes (R. M. Nilsen et al., 2009). However, it is important to note that the nature and severity of selection bias depend on the specific exposure-outcome association being examined and the dataset used. Therefore, results from studies of selection bias cannot be generalised to make broad assumptions about the overall nature of selection bias in cohort studies that include multiple exposures and outcomes (Hernán et al., 2004). For example, research using data from the Avon Longitudinal Study of Parents and Children (ALSPAC; Fraser et al., 2013) showed that selective attrition can result in underestimation of socioeconomic inequalities in various health-related outcomes, and the degree of bias worsens as the proportion of drop-outs increases (Howe et al., 2013). This highlights that that the impact of selection bias can vary depending on the specific exposure-outcome association being examined and the dataset used. Thus, it is important to consider the specific context and characteristics of each study when assessing the potential impact of selection bias.

Recently, some genetic association studies have sought to identify factors influencing selective participation to improve understanding of selection biases in such studies. Genome-wide association studies (GWAS) have found genome-wide significant evidence for a number of loci associated with continued participation in several cohorts including ALSPAC, UK Biobank and the Norwegian Mother, Father and Child Birth Cohort (MoBa; Adams et al., 2020; Ask et al., 2021; Taylor et al., 2018; Tyrrell et al., 2020). In a study of the ALSPAC cohort, common genetic variants were estimated to explain 18-32% of variability in continued participation phenotypes measured by the number of questionnaires/clinics completed and completion of the most recent clinic/questionnaire for both mother and child (Taylor et al., 2018). In UK Biobank, SNP-based heritability of continued participation in a follow-up mental health questionnaire was estimated to be 9.9% (Adams et al., 2020). In MoBa, SNP-based heritability of participation in the most recent data collection (when children were 8 years old) was

estimated to be 6.5% for mothers and 14.7% for fathers, and SNP-based heritability of continuously measured participation in mothers (number of questionnaires returned) was estimated to be 17% (Ask et al., in prep). These estimates of SNP-based heritability highlight the potential influence of genetic factors on selective participation. By understanding the genetic underpinnings of continued participation, researchers may gain further insights into the mechanisms driving selection biases in cohort studies and help contribute to the development of more refined and sophisticated methods to address selection bias.

Studies have also demonstrated genetic overlap between continued participation and several traits. In ALSPAC, continued participation was associated with polygenic scores for higher educational attainment, agreeableness and openness, whereas attrition was associated with higher polygenic scores for BMI, smoking initiation, neuroticism, schizophrenia, ADHD and depression (Taylor et al., 2018). A study in UK Biobank showed that follow-up participation in a mental health questionnaire was genetically correlated with better health, higher educational attainment, and lower rates of psychiatric disorders (Adams et al., 2020). Tyrrell et al. (2020) used data from UK Biobank to explore genetic correlates of continued participation in four optional surveys. Results were consistent with previous studies showing positive genetic correlations between participation and academic qualifications, fluid intelligence and educational duration, and inverse genetic correlations between participation and obesity-related traits. Mendelian randomisation analyses further showed that longer educational duration, later age at menarche and height increased participation whereas obesity, dyslipidaemia, Alzheimer's disease, neuroticism and schizophrenia reduced participation in the optional components. Strong positive genetic correlations were also found between continued participation in ALSPAC and three of the optional UK Biobank surveys indicating that similar genetic factors are associated with continued participation and follow-up in optional surveys across the two studies.

Studies indicate that selection bias can impact upon the magnitude of associations between polygenic scores and phenotypic outcomes. For example, Munafò et al., (2018) compared association estimates in a selected sub-study of ALSPAC (a sub-set of mother-offspring pairs who were selected based on availability of DNA samples at multiple time points in both generations; Relton et al., 2015) with estimates derived from the full cohort and found that

polygenic risk for smoking was associated with maternal education in the selected sub-sample but not in the full sample. This would suggest that conditioning on selection (i.e., by analysing data only within a selected sub-sample) can induce associations between polygenic score and outcome that are not present in the wider population. Findings from another study using ALSPAC data showed that associations between polygenic risk for education and being an ever smoker, and between the education polygenic score and BMI, were attenuated in a subsample of those who attended the most recent clinic compared with the full genetic sample (Taylor et al., 2018). These findings illustrate the potential to introduce bias into genetic analyses when studying selected subsamples based on the availability of follow-up data.

Overall, past research has shown that it is possible for selective participation to bias associations between factors associated with participation in a study. Selection bias has been demonstrated to influence estimated associations between a range of sociodemographic and health-related characteristics, and to influence estimated genetic risk linked to these characteristics. To date, most studies focussed on genetic influence in selection bias have focussed on polygenic scores or SNP-based heritability. These are known to capture only a portion of total heritability (Maher, 2008), so it is unclear the extent to which genetic factors contribute to study participation. To our knowledge, no studies have used a behavioural genetic approach to estimate the heritability of continued participation, nor genetic and environmental overlap with factors associated with continuing participation.

In the present study, our aim was to complement and extend previous research investigating the influence of selection bias on phenotypic and genetic associations in the context of a population-based cohort study. Specifically, we sought to examine factors associated with continued participation in the Norwegian Mother, Father and Child Birth Cohort Study (MoBa). We investigated associations between maternal participation and various baseline variables, socioeconomic factors, and maternal and offspring outcomes. Additionally, we explored potential bias due to selection on associations between maternal educational attainment or income rank and outcomes related to maternal and offspring internalising and externalising symptoms. Further, we employed biometric model fitting to estimate genetic and environmental influences associated with maternal continued participation, educational attainment, and internalising symptoms. We next examined genetic and environmental

correlations between maternal educational attainment, internalising symptoms, and continued participation using extended bivariate twin models. Lastly, we explored potential bias due to selection on genetic and environmental correlations between maternal educational attainment and internalising symptoms. To do so, we used extended bivariate models to examine the relationship between these variables in both the full sample and among the subset of individuals who participated at the 8-year follow-up assessment in MoBa.

4.3. Methods

4.3.1. Sample

Data were drawn from Norwegian administrative registers provided by Statistics Norway and from the Norwegian Mother, Father and Child Birth Cohort Study (MoBa) (Magnus et al., 2016). MoBa is a prospective population-based pregnancy cohort study conducted by the Norwegian Institute of Public Health. Recruitment into MoBa occurred from 1999 to 2009 at 50 of Norway's 52 hospitals during routine ultrasound examinations offered to all pregnant women at gestational weeks 17-18. Invitations were issued to women in 277,702 pregnancies, with an initial participation rate of 41%. The initially consenting mothers were invited to complete eight follow-up questionnaires on themselves and their children, and two questionnaires were sent to fathers. The cohort includes ~114,500 children, 95,200 mothers, and 75,00 fathers who have participated in at least one wave of data collection. Kinship between MoBa participants has been identified through linkage with pedigree and zygosity information from the Medical Birth Registry of Norway and the Norwegian Twin Registry, respectively (NTR) (T. S. Nilsen et al., 2013). The current sample comprised a total of 54,763 mothers. We grouped MoBa participants into extended families, identified via pairs of siblings (twins, full-siblings, or half-siblings) or first cousins in the parent generation. Mothers who did not have participating extended family members were included in phenotypic analyses only. Phenotypic data were drawn from version 12 of the quality assured MoBa datafiles. In total, raw phenotypic data were used for 54,763 mothers. Table S1 shows frequencies of mothers stratified by relatedness group. The current study also uses national register data on mothers' education and income. The Norwegian system of personal identification numbers was used to link register data with MoBa data.

4.3.2. Ethics

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 now based on regulations related to the Norwegian Health Registry Act. The current study was approved by The Regional Committees for Medical and Health Research Ethics.

4.3.3. Measures

4.3.3.1. Continued participation in MoBa

Participants were invited to fill out questionnaires three times during pregnancy (15, 22, and 30th weeks of gestation) and five times after birth (when children were 6 months, and 1.5, 3, 5 and 8 years old). Some of the data collections were delayed so mothers of the oldest children were not invited to all the assessments (Magnus et al., 2016). Maternal continued participation across all data collections was measured by counting the total number of questionnaires returned. When divided by the number of assessments invited to, this gave a continuous variable indicating the proportion of questionnaires returned among the invited mothers and was scaled between 0 and 1.

4.3.3.2. Participation predictors

Baseline demographic variables

Baseline demographic variables potentially associated with participation were extracted from the Medical Birth Registry of Norway (MBRN) data. These variables were: mothers age at birth, parity (i.e., number of births), offspring year of birth, and offspring sex.

Socioeconomic factors

We obtained education and income data from national registries. The data included information on mothers' educational attainment and total income (the sum of income from work capital gains and benefits received during the calendar year) from 1999 to 2009. These

years correspond to the birth years of the first and last child in the MoBa cohort. We analysed mothers' education and income corresponding to the year they gave birth to their first child in MoBa. Level of educational attainment was indexed in accordance with the Norwegian Standard Classification of Education (Statistics Norway, 2013), with values ranging from 1 (primary education) to 8 (doctoral-level education). Regarding mothers' income, we created a measure of income rank that indicates an individual's position in the distribution of incomes within a cohort-sex-year-group (e.g., income for females born in 1983). The income rank was scaled between 0 and 1, with higher values denoting a higher income rank within the reference group.

Maternal internalising symptoms

Mothers' internalising symptoms were measured by self-report at 15 weeks of gestation, using the five-item version of the short form of the Hopkins Symptom Checklist (SCL-5; Tambs & Moum, 1993). The scale consists of five statements relating to symptoms of anxiety (e.g., nervousness or shakiness) and depression (e.g., feeling hopeless about the future) experienced in the last two weeks, scored on a scale from 1 (not at all) to 4 (very much). Scores were calculated as the average of the five items, with higher scores indicating greater symptoms of anxiety and depression. Participants who were missing data on more than half of the items were considered missing.

Offspring internalising and externalising symptoms

Offspring internalising and externalising symptoms were measured by maternal report when children were 18 months old, using items from the Child Behaviour Checklist (CBCL) for preschool children (Achenbach, 1992). The internalising sub-scale included five items that measure emotional symptoms, and the externalising sub-scale included eight items that measure behavioural symptoms. Mothers reported agreement for each item based on a three-point Likert scale: 1 = Not true; 2 = Somewhat true; 3 = Very/often true. Mean scores for internalising and externalising symptoms were calculated, excluding participants missing on more than half of the items.

4.3.4. Statistical analysis

We conducted a series of analyses to obtain the following estimates: (1) associations between participation predictors and continued participation; (2) moderation of associations between mothers' socioeconomic factors and internalising symptoms, and offspring internalising and externalising symptoms; (3) estimates from univariate heritability analyses; (4) estimates from bivariate heritability analyses investigating the relationships between maternal educational attainment, internalising symptoms and continued participation. All analyses were carried out in R v.4.0.3 (https://www.R-project.org/; R Core Team, 2020).

4.3.4.1. Phenotypic associations

Predictors of continued participation

Linear regressions were performed to assess associations between participation predictors (i.e., maternal baseline demographic variables, maternal socioeconomic factors, maternal internalising symptoms, offspring internalising and externalising symptoms) and continued participation.

Impact of selective participation on exposure-outcome associations

To investigate the impact of selective participation on exposure-outcome estimates, we conducted moderation analyses using linear regression to examine whether continued participation moderated associations between maternal socioeconomic factors (i.e., educational attainment and income rank) and the following outcomes: maternal internalising symptoms, offspring internalising symptoms, and offspring externalising symptoms.

To model possible variation in the effect of maternal education on these outcomes for different values of continued participation, we included continued participation as an additive interaction term in linear regressions, with maternal educational attainment as the predictor and (1) maternal internalising scores as the outcome variable; (2) offspring internalising scores as the outcome variable; (3) offspring externalising scores as the outcome variable.

To evaluate whether attrition may bias the estimated effect of maternal income on the outcome variables for different values of continued participation, we included continued participation as an additive interaction term in linear regressions, with maternal income rank as the predictor and (1) maternal internalising scores as the outcome variable; (2) offspring internalising scores as the outcome variable; (3) offspring externalising scores as the outcome variable.

4.3.4.2. Biometric model fitting

The classical twin design compares intra-class correlations of identical (monozygotic, MZ) and non-identical (dizygotic, DZ) twins to estimate the contribution of genetic and environmental factors to observed phenotypic variance in a trait and/or covariance between traits (Rijsdijk & Sham, 2002). The classical twin method is based on the following assumptions: (1) MZ twins share 100% of their genes and DZ twins share on average 50% of their segregating genes; (2) MZ and DZ twin pairs share environmental influences common to both twins in the same family to the same extent ('shared environment'); and (3) MZ and DZ twins differ from one another due to exposure to environmental factors which are unique to the individual ('non-shared environment'). The similarity of reared-together twins is attributed to additive genetic (A) and shared environmental (C) factors that are common to both twins in the same family. The correlation between twins' shared environment is assumed to be 1 for both MZ and DZ pairs. The differences between MZ and DZ twin pairs is attributed to non-shared environmental factors (E), which are unique to the individual. We adapted these twin models to account for the additional degrees of genetic relatedness between siblings (50%), half-siblings (25%) and cousins (12.5%), and to constrain the shared environmental effect in cousins and paternal half-siblings to zero (as most cousins and paternal half-siblings do not share a household).

Structural equation model-fitting analyses were performed to estimate the relative contribution of additive genetic (A), shared environment (C), and non-shared environment (E) factors to the variation in maternal continued participation, educational attainment, and internalising symptoms. To test for the significance of the A and C components to the variance of the traits, we compared each full model with a model in which the C parameter was

dropped (AE model) and with a model in which the A and C parameters were dropped (E model). This technique further extends to bivariate analyses, in which the variance in two phenotypes, and the covariance between them, is partitioned into genetic, shared environmental, and non-shared environmental effects. The extent to which genetic and environmental influences are correlated between phenotypes is calculated by estimating genetic correlations (rA), shared environmental correlations (rC) and non-shared environmental correlations (rE). Extended bivariate twin models were applied to assess genetic and environmental relationships between: (1) maternal educational attainment and continued participation; and (2) maternal internalising symptoms and continued participation. To investigate the impact of selective participation at follow-up on estimates of genetic and environmental correlations, extended bivariate analyses were carried out to assess genetic and environmental relationships between exposure-outcomes with significant phenotypic moderation (in this case, between maternal educational attainment and internalising symptoms) among (i) all individuals included in the present sample and (ii) among individuals who returned the latest MoBa questionnaire included in our analysis administered when children were 8 years old. We tested the significance of the rA and rC correlations by comparing the relative fit of the full model with models in which each parameter was dropped. The full univariate and bivariate model specifications are shown in Supplementary Figures S1 and Figure S2, respectively.

All SEM analyses were conducted using the open-source package OpenMx (Neale et al., 2016). OpenMx uses full-information maximum-likelihood to estimate model parameters. Prior to model fitting, study variables were regressed on the following covariates: maternal age, parity (number of births), and child year of birth, and standardised residuals were used. The residual score for internalising symptoms was log transformed to reduce positive skew. Models were fitted to raw data using full-information maximum-likelihood estimation and compared using likelihood ratio testing (differences in -2 Log-likelihood estimation and degrees of freedom, which is c^2 distributed) and the Akaike's Information Criterion (AIC) (Akaike, 1987).

4.4. Results

4.4.1. Descriptive statistics

Table 1 presents an overview of the study sample. Participation rates were highest in pregnancy and showed a gradual decline over time, particularly when the children reached early childhood (Figure 1; Table S2). Mothers at recruitment were aged on average 29.8 (SD = 4.34) years. The percentages of women in the study sample who had previously 0, 1, 2, or 3 or more children are 45.7%, 37.3%, 13.5%, 3.5%, respectively. 51% of the children born to participating mothers were males. Mothers with higher education attainment (undergraduate or postgraduate level) constituted 64.6% of the MoBa sample (M = 5.32, SD = 1.42), compared to around 46% of the target population. The mean income rank among participating mothers was 0.60 (SD = 0.25), indicating that, on average, the sample had a slightly higher income compared to the general population. Total scores for mothers' internalising symptoms were low on average (M = 1.27, SD = 1.27) compared to population norms (Strand et al., 2009), showing positively skewed distributions (Figure S3). Total scores for offspring internalising (M = 1.26, SD = 0.24) and externalising (M = 1.49, SD = 0.28) symptoms were low on average and showed positively skewed distributions.

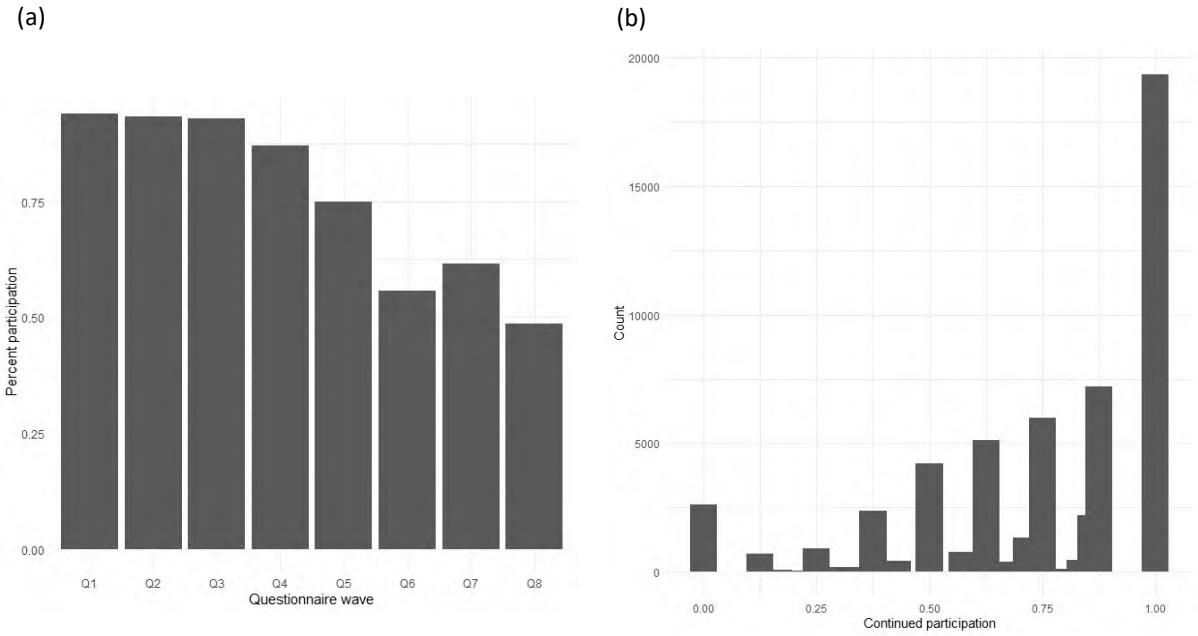


Figure 1. Mothers' participation rates in the Norwegian Mother, Father and Child Birth Cohort Study (MoBa). Figure 1a shows the percentage of participation among invited mothers at each wave of assessment. Figure 1b depicts the frequency distribution of continued participation indicating the proportion of questionnaires returned among the invited mothers.

4.4.2. Phenotypic associations

Associations between maternal participation and baseline variables (mothers age at birth, parity, year of birth and child sex), socioeconomic factors (mothers' educational attainment and income rank) and internalising symptoms are shown in Table 1. Older age at birth, having no previous births (i.e., mothers of firstborn children), and early recruitment/birth of index child was associated with higher rates of participation. Higher educational attainment was associated with continued participation, indicating that mothers with higher educational attainment when their child was born were more likely to continue participating over time. Higher maternal internalising scores were negatively associated with continued participation, indicating that mothers who self-reported higher internalising symptoms in the first questionnaire (at 15 weeks of gestation) were less likely to continue to participate over time. There was a small association between mothers' income rank and continued participation, however, this effect became non-significant after adjusting for all other variables included in the analysis. Offspring internalising and externalising scores were not significantly associated with mothers' study participation over time.

Table 1. Associations between participation predictors and continued participation

	Total (N=54763)	**Continued participation Unadjusted b**	**Continued participation Adjusted b**
Maternal age (at birth of index child)			
Mean (SD)	29.8 (4.34)	0.06 (0.05, 0.07)	0.03 (0.02, 0.03)
Missing	17 (0.0%)		
Parity (number of births)			
0 (ref)	25007 (45.7%)	-	-
1	20422 (37.3%)	-0.17 (-0.19, -0.15)	-0.03 (-0.04, -0.02)
2	7416 (13.5%)	-0.21 (-0.24, -0.18)	-0.06 (-0.07, -0.04)
3 or more	1918 (3.5%)	-0.30 (-0.35, -0.26)	-0.07 (-0.10, -0.04)
Missing	0		
Child year of birth			
Mean (SD)	2010 (1.92)	-0.12 (-0.13, -0.11)	-0.06 (-0.06, -0.05)
Missing	0		
Child sex			
Male (ref)	28158 (51.4%)		-

Female	26509 (48.4%)	0.01 (0.00, 0.03)	0.00 (-0.01, 0.01)
Missing	96 (0.2%)		
Educational attainment			
Mean (SD)	5.32 (1.42)	0.18 (0.18, 0.19)	0.06 (0.06, 0.07)
Missing	1128 (2.1%)		
Income rank			
Mean (SD)	0.60 (0.25)	0.06 (0.05, 0.07)	-0.01 (-0.01, 0.00)
Missing	970 (1.8%)		
Maternal internalising symptoms			
Mean (SD)	1.25 (0.40)	-0.07 (-0.08, -0.06)	-0.02 (-0.03, -0.02)
Missing	3816 (7.0%)		
Offspring internalising symptoms			
Mean (SD)	1.26 (0.24)	-0.01 (-0.02, -0.01)	0.00 (0.00, 0.01)
Missing	16457 (30.1%)		
Offspring externalising symptoms			
Mean (SD)	1.49 (0.28)	-0.01 (-0.02, -0.01)	-0.01 (-0.01, 0.00)
Missing	15970 (29.2%)		

Note. Linear regressions were conducted using standardised outcome variables. Adjusted B coefficients were calculated after including all other variables in the table.
*p < 0.05; **p < 0.01

4.4.3. Impact of selective participation on phenotypic exposure-outcome associations

We conducted moderation analyses to assess potential bias due to selection on association estimates between mothers' educational attainment or income rank with the following outcomes: mothers' internalising symptoms, offspring internalising symptoms and offspring externalising symptoms.

<u>Outcome: Mothers' internalising symptoms</u>

Educational attainment. Results showed a significant main effect of maternal educational attainment on internalising symptoms (B = -0.11; 95% CI [-0.12, -0.10]), and a main effect of continued participation on internalising symptoms (B = -0.06, 95% CI = [-0.07, -0.05]). We

found a significant interaction term between educational attainment and continued participation (B = 0.03; 95% CI [0.02, 0.04]). This indicates that the association between educational attainment and internalising symptoms varies across different levels of participation (Figure 2). To follow-up this interaction, we calculated the estimated marginal means (EMMs) of internalising symptoms at different levels of education (-1 SD/mean/+1 SD) across different levels of participation (-1 SD/mean/+1 SD) and tested for the effect of educational attainment at different levels of participation with simple slope analysis. The EMMs of internalising symptoms at different levels of education across different levels of participation are shown in Table S1. Results from the simple slope analysis showed that the association between educational attainment and internalising symptoms was attenuated among those with above average participation (+1 SD) (B = -0.08; 95% CI [-0.09, -0.07]), and among those with average participation (B = -0.11; 95% CI [-0.12, -0.10]), compared to those with below average (-1 SD) participation (B = -0.14; 95% CI [-0.15, -0.12]). This indicates that the association between educational attainment and internalising symptoms was weaker at higher levels of continued participation.

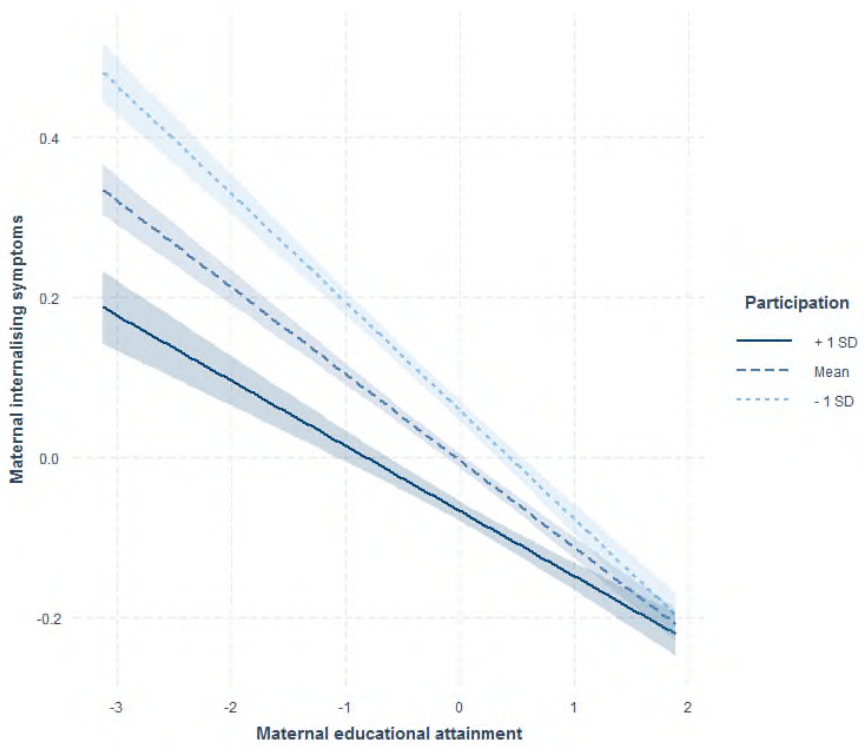


Figure 2. Interaction plot depicting the relationship between maternal educational attainment and internalising symptoms across the different levels (+1 SD, mean, -1 SD) of continued participation.

Income rank. Higher internalising scores among mothers was associated with lower income rank (B = -0.07; 95% CI [-0.08, -0.06]) and lower rates of continued participation (B = -0.08; 95% CI [-0.08, -0.06]). No significant moderation effects of continued participation on the association between income rank and internalising symptoms were found (B = 0.01; 95% CI [0.00, 0.02]). This suggests minimal variation in the association between mothers' income rank and internalising symptoms as a function of varying rates of continued participation.

Outcome: Offspring internalising symptoms

Educational attainment. Lower maternal educational attainment was associated with higher internalising scores among offspring (B = -0.10; 95% CI [-0.11, -0.09]. Continued participation was not significantly associated with offspring internalising symptoms in our analyses (B = -0.00, 95% CI [-0.01, 0.01]). No moderation effects of continued participation on the association between mothers' educational attainment and offspring internalising symptoms were found (B = 0.01, 95% CI [0.00, 0.02]).

Income rank. Higher offspring internalising symptoms was associated with lower maternal income rank (B = -0.05, 95% CI [-0.06, -0.04]) and lower continued participation (B = -0.02; 95% CI [-0.03, -0.01]). No significant moderation effects of continued participation on the association between mothers' income rank and offspring internalising symptoms were found (B = 0.00, 95% CI [-0.01, 0.01).

Outcome: Offspring externalising symptoms

Educational attainment. Higher offspring externalising scores was associated with lower maternal educational attainment (B = -0.08, 95% CI [-0.09, -0.07]) and lower rates of continued participation (B = -0.01, 95% CI [-0.03, -0.01]). The association between mothers' educational attainment and offspring externalising symptoms was not found to be moderated by participation (B = 0.01; 95% CI [0.00, 0.02]).

Income rank. Higher offspring externalising scores was associated with lower maternal income rank (B = -0.02, 95% CI [-0.03, -0.01]) and lower rates of continued participation (-0.02, 95%

CI [-0.03, -0.01]). No significant moderation effects of participation on the association between mothers' income rank and offspring externalising symptoms were found (B = 0.01; 95% CI [0.00, 0.02]).

4.4.4. Biometric model fitting

4.4.4.1. Genetic and environmental influences associated with maternal continued participation, educational attainment, and internalising symptoms

ACE parameter estimates from univariate models are presented in Table 2. For all phenotypes, C did not significantly account for measure variance, so were subsequently dropped. Exclusion of C resulted in a more parsimonious model for all phenotypes, without significant changes to model fit (Table S2). Genetic influences explained 34% (95% CI [0.27, 0.40]) of the variance in maternal continued participation and non-shared environmental influences explained 67% (95% CI [0.59, 0.73]) of the variance. Genetic influences explained 75% (95% CI [0.70, 0.81]) of the variance in maternal educational attainment and non-shared environmental influences explained 25% (95% CI [0.19, 0.30]) of the variance. Variance in maternal internalising symptoms was explained by genetic influences (31%; 95% CI [0.24, 0.39]) and non-shared environmental influences (69%; 95% CI [0.61, 0.76]).

4.4.4.2. Relationship between maternal educational attainment, internalising symptoms, and continued participation

Two extended bivariate models were conducted to examine genetic and environmental correlations between mothers' continued participation and (1) their educational attainment, or (2) internalising symptoms. Results from the best-fitting models are depicted in Figure 3. Fit statistics from bivariate model-fitting analyses are shown in Supplementary Table S3.

Maternal educational attainment – continued participation

We observed a moderate phenotypic correlation between maternal educational attainment and continued participation (rPh = 0.17, 95% CI [0.16, 0.18]). Genetic influences on educational attainment were significantly correlated with genetic influences on continued participation (*Ra* = 0.33; 95% CI [0.18, 0.49]). Shared environmental factors between continued participation and educational attainment were also correlated (*Rc* = 0.62; 95% CI

[0.39, 1.00]). Non-shared environmental factors did not significantly contribute to the covariance of these traits.

Maternal internalising symptoms – continued participation

Maternal internalising scores showed a negative phenotypic correlation with continued participation (rPh = -0.11; 95% CI [-0.12, -0.09]). Genetic influences on internalising symptoms were significantly correlated with genetic influences on continued participation (*Ra* = -0.41, 95% CI [-0.63, -0.22]). No significant shared environmental or non-shared environmental correlations were found.

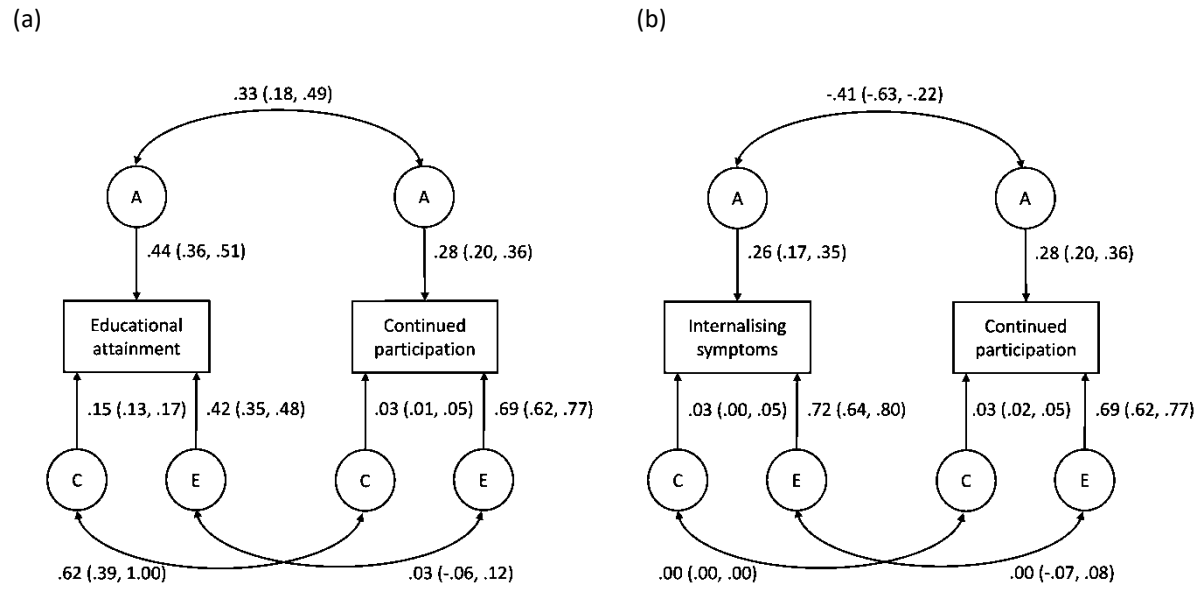


Figure 3. Extended bivariate model fitting results showing the relationships between (a) educational attainment and continued participation and (b) internalising symptoms and continued participation. Note. A = additive genetic influences, C = shared environmental influences, E = non-shared environmental influences

4.4.4.3. Impact of selective participation on genetic and environmental correlations between maternal educational attainment and internalising symptoms

Two extended bivariate models were conducted to examine genetic and environmental correlations between maternal educational attainment and internalising symptoms among (i)

all individuals and (ii) individuals who returned the MoBa questionnaire at 8 years. Results from the best-fitting models for the full sample and for the subset who returned the questionnaire at 8 years are depicted in Figure 4. Fit statistics from bivariate model-fitting analyses are shown in Supplementary Table S4.

(i) *Full sample.* Maternal educational attainment showed a negative phenotypic correlation with internalising scores (rPh = -0.12, 95% CI [-0.13, -0.11]). Genetic influences on educational attainment were significantly correlated with genetic influences on internalising symptoms (*Ra* = -0.40, 95% CI [-0.57, -0.25]). No significant shared environmental or non-shared environmental correlations were found.

(ii) *Subsample.* The phenotypic correlation between maternal educational attainment and internalising scores in the subsample (rPh = -0.10, 95% CI [-0.11, -0.08]) was of similar magnitude to that observed in the full sample. Genetic influences on educational attainment were significantly correlated with genetic influences on internalising symptoms (*Ra* = -0.29, 95% CI [-0.58, -0.04]). Overlapping confidence intervals indicated that the genetic correlation was of similar magnitude to that observed in the full sample. Consistent with the results from the full sample, no significant shared environmental or non-shared environmental correlations between maternal educational attainment and internalising symptoms were found.

4.5. Discussion

The present study aimed to evaluate bias due to attrition in the MoBa cohort and examine the potential impact of selective participation on phenotypic and genetic exposure-outcome associations using an extended family design. The results provide insights into the characteristics of participating mothers in MoBa and the influence of selective participation on various outcomes.

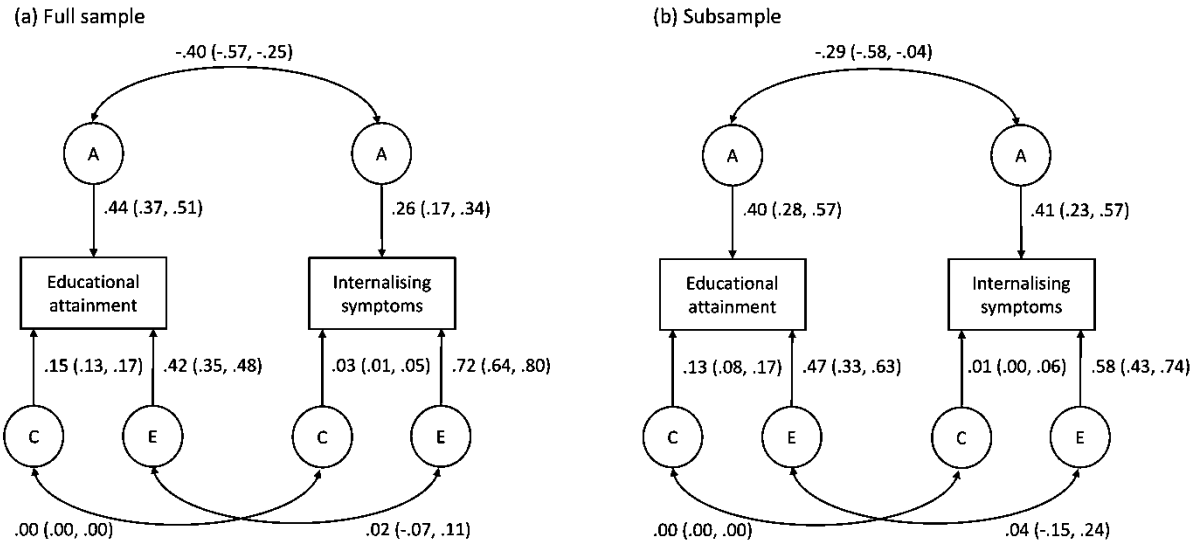


Figure 4. Extended bivariate model fitting results showing the relationship between maternal educational attainment and internalising symptoms from (a) the full sample and (b) the subsample of individuals who returned the MoBa questionnaire at 8 years. Note. A = additive genetic influences, C = shared environmental influences, E = non-shared environmental influences.

4.5.1. Factors associated with continued participation

Participation rates were highest during pregnancy and gradually declined over time, particularly when the children reached early childhood, and a substantial proportion of the sample had no previous children (45.7%). A higher proportion of mothers in the study sample had higher education attainment (undergraduate or postgraduate level) compared to the target population (64.6% versus ~46%) and the sample had a slightly higher income rank compared to the general population (0.60 vs.0.50). Further, our findings indicated that social advantage was associated with higher rates of continued participation in MoBa. Specifically, we observed that older mothers (based on age at the birth of the index child), mothers without previous children before initial participation, and mothers with higher educational attainment were more likely to continue participating. There was a small association between higher income rank and continued participation, however, this became non-significant after adjusting for other variables. Conversely, later recruitment and higher levels of internalising symptoms were associated with attrition over time. These findings are similar to those reported in previous studies conducted in MoBa (Biele et al., 2019; Vejrup et al., 2022), as well as in other prospective birth cohorts (Bliddal et al., 2018; Cornish et al., 2020) and population-

based cohorts (Tyrrell et al., 2021). Taken together, these findings suggest that attrition in cohort studies often leads to an overrepresentation of individuals who are more advantaged in terms of health and socioeconomic status than those who drop-out.

4.5.2. Impact of selective participation on phenotypic exposure-outcome associations

We explored the potential bias introduced by selective participation on observed associations between maternal educational attainment or income rank and internalising (mothers' and offspring) and externalising (offspring only) symptoms. For maternal internalising symptoms, we found a significant interaction effect between educational attainment and continued participation, indicating that the association between educational attainment and internalising symptoms varied depending on the level of continued participation. In line with previous findings (Howe et al., 2013), we found that the association between educational attainment and internalising symptoms was weaker among those with above-average and average participation rates compared to those with average or below-average participation. The significant interaction effect between maternal educational attainment and continued participation in relation to internalising symptoms raises the possibility of collider bias (Munafò et al., 2018). Collider bias occurs when conditioning on a variable, in this case, continued participation, that is influenced by both the predictor (educational attainment) and the outcome (internalising symptoms). The weaker association observed among those with above-average and average participation rates suggests that conditioning on continued participation, a potential collider, may have introduced bias into the estimate of the association between educational attainment and internalising symptoms. This suggests bias could compromise the internal validity of the associations and highlights the potential impact of collider bias on the study findings. Such bias could be problematic particularly if similar characteristics related to participation in different studies as the results might be reproducible but may equally be biased. This is an issue for research combining results from observational studies as if results from such studies are biased, combing these (biased) estimates will result in an overall precise yet biased finding. Additionally, errors in the estimation of exposure-outcome associations can be problematic for translational research (Sullivan & Feinn, 2012). However, our results suggest that the nature and extent of bias is specific to the exposure-outcome associations being tested. Specifically, no moderation effects of continued

participation were observed for the association between mothers' income rank and internalising symptoms, nor for the associations between mothers' socioeconomic factors and offspring outcomes. Thus, differences between factors associated with participation across studies may partly explain heterogeneity between studies. Identifying predictors of participation informs researchers about potential bias and aids in choosing appropriate statistical approaches to reduce the impact of bias due to attrition (e.g., inverse probability of participation weighting [IPPW]; (Metten et al., 2022)). Future research could consider investigating additional mediators and moderators to comprehensively understand the complex interplay between participation, socioeconomic factors, and mental health outcomes.

4.5.3. Aetiology of continued participation

Genetic influences explained a substantial proportion of the variance in maternal continued participation (34%), while non-shared environmental influences accounted for the remaining variance. Our finding of genetic influences on continued participation in MoBa aligns with recent research indicating a genetic contribution to participation phenotypes. Previous studies using genotype data have found evidence of genetic variants associated with ongoing study participation and reported SNP-based heritability estimates of participation phenotypes ranging from 6.5% to 27% Adams et al., 2020; Ask et al., 2021; Taylor et al., 2018; Tyrrell et al., 2020). Observing genetic influences on continued participation suggests the presence of gene-environment correlation because it indicates that genetic factors are associated with both the inclination to participate in longitudinal research and the likelihood of sustained engagement over time. Gene-environment correlation refers to the concept that genetic factors can influence an individual's exposure to certain environments or experiences. In this case, the genetic factors that influence continued participation in the MoBa study are likely associated with various traits or characteristics that impact an individual's willingness to participate in research.

4.5.4. Relationship between maternal educational attainment, internalising symptoms, and continued participation

4.5.4.1. Educational attainment – continued participation

We observed a moderate phenotypic association between mothers' educational attainment and continued participation, such that mothers with higher educational attainment were more likely to continue to participate at follow-up. Genetic factors related to educational attainment were also related to continued participation in MoBa. This is consistent with previous studies demonstrating that polygenic scores for educational attainment predict continued participation at follow-up assessments (Adams et al., 2020; Taylor et al., 2018). The observed genetic correlation between educational attainment and continued participation suggests that the genetic factors influencing these two variables were partially shared. This suggests that there are likely genetic factors that contribute to both higher educational attainment and increased participation in longitudinal studies like MoBa. One possible explanation for this genetic overlap is the presence of certain genetically influenced traits or characteristics that promote sustained engagement in study participation over time. These traits may manifest in multiple domains, including educational pursuits and willingness or motivation to participate in research studies. For example, genetic factors influencing cognitive abilities, personality traits (e.g., openness to new experiences), or socioemotional characteristics, may contribute to both higher educational attainment and continued participation in cohort studies. This aligns with prior evidence of genetic overlap between continued participation and traits such as agreeableness, openness, and fluid intelligence (Adams et al., 2020; Taylor et al., 2018).

Our results also showed that shared environmental factors were correlated between mothers' educational attainment and continued participation. This suggests the presence of shared environmental influences that specifically contribute to the covariation between educational attainment and continued participation. A possible explanation could be that shared environmental factors that promote educational opportunities and access to resources may also facilitate participation in research studies.

Notably, non-shared environmental factors did not significantly contribute to the covariance of these traits. In the context of twin or family studies, the absence of a non-shared environmental correlation is generally interpreted as evidence against a causal relationship between variables (McAdams et al., 2021). Non-shared environmental factors refer to unique experiences or circumstances that differ between individuals within the same family. If there is a causal relationship between two variables, it would be expected that individual-specific environmental factors contribute to their relationship. Thus, the absence of a non-shared environmental correlation suggests a non-causal relationship between educational attainment and continued participation.

4.5.4.2. Internalising symptoms – continued participation

Turning to the relationship between mothers' internalising symptoms and continued participation, we observed a negative phenotypic correlation between these variables, indicating that mothers who reported higher internalising symptoms were less likely to continue to participate over time. Results showed that genetic influences on these traits were correlated. This is consistent with previous findings of an association between polygenic risk for depression and continued participation at follow-up assessments (Adams et al., 2020; Taylor et al., 2018). The significant genetic correlation observed between mothers' internalising symptoms and continued participation suggests that genetic factors contributing to internalising symptoms may also influence the likelihood of continued participation. One possible explanation is genetic pleiotropy, wherein shared genetic variants contribute to both internalising symptoms and continued participation but have opposite effects on these traits. Certain genetic factors may increase susceptibility to internalising symptoms while simultaneously reducing the likelihood of continued participation. This pleiotropic effect could account for the negative genetic correlation observed in our study.

4.5.5. Impact of selective participation on the relationship between maternal educational attainment and internalising symptoms

To explore the impact of selection bias on the relationship between mothers' educational attainment and internalising symptoms in MoBa, we investigated the genetic and environmental correlations between these variables in the full sample and in the subset of

individuals who returned the MoBa questionnaire at 8 years. Analysis in the full sample showed that genetic influences on maternal educational attainment were correlated with genetic influences on (maternal) internalising symptoms. The genetic correlation in the sub-sample was of similar magnitude to that observed in the full sample, suggesting minimal impact of selection bias on the observed genetic correlation between maternal educational attainment and internalising symptoms. Past studies investigating the impact of selection bias on associations between polygenic scores and various outcomes have shown that attrition may bias some, but not all, association estimates. For example, one study in ALSPAC showed that attrition attenuated associations between the polygenic score for education and smoking or BMI, whereas associations between polygenic scores for BMI, smoking and schizophrenia and self-reported BMI and smoking were of similar magnitude to those observed in the full genetic sample (i.e., the total sample for which DNA data is available) (Taylor et al., 2018). These findings illustrate the potential to introduce bias in some, though not most, estimates of genetic associations when studying selected subsamples based on the availability of follow-up data.

4.5.6. Limitations

It is important to acknowledge some limitations of this study. First, our study specifically focussed on the potential for bias due to selection mechanisms at follow-up assessment, rather than bias associated with selection at baseline (i.e., bias associated with initial recruitment). Future research with additional linkage to registry data on relevant exposure and outcome characteristics for the entire target population should seek to investigate biases attributable to differences between participants and non-participants. Second, our results may not be generalisable to studies with different selection criteria or recruitment methods, or to studies with specific cultural or contextual factors that may influence study participation. Factors influencing participation may also change over time and/or with age. Third, we did not explore all possible traits that might be associated with participation in the MoBa cohort. Future studies should evaluate the impact of selection bias on associations between other socioeconomic factors and mental health outcomes.

4.6. Conclusion

Overall, this study highlights the factors associated with mothers' continued participation in the MoBa cohort and the potential impact of selective attrition on phenotypic exposure-outcome associations. Our findings revealed that certain characteristics, such as older age at birth, being a first-time mother, and higher educational attainment, were associated with higher rates of continued participation. Selective participation influenced the observed associations between maternal educational attainment and internalising symptoms, with associations being weaker among individuals with average or below-average participation rates. However, no moderation effects of continued participation were observed for the association between maternal income rank and internalising symptoms, or for the associations between maternal socioeconomic factors and offspring outcomes. Genetic factors played a substantial role in continued participation, suggesting the presence of gene-environment correlation. Our study also demonstrated the presence of shared genetic factors that contribute to the covariation between maternal educational attainment or internalising symptoms and continued participation. The impact of selection bias on the relationship between maternal educational attainment and internalising symptoms was minimal, as the observed genetic correlation remained similar in the full sample and the subset of individuals with follow-up data at 8 years. Nonetheless, selection bias can distort the observed associations between certain exposure and outcome variables, and caution should be exercised when generalising findings from samples characterised by selection to the larger population.

Identifying predictors of participation informs strategies to mitigate bias, such as inverse probability of participation weighting or imputation methods (Metten et al., 2022). Further, researchers could consider strategies to improve study participation and representation at study inception. Patient and Public Involvement and Engagement (PPIE) has the potential to assist researchers, healthcare professionals, and stakeholders in effectively planning and conducting research studies, as well as implementing health policies, that better align with the specific requirements of the community (National Institute for Health Research, 2023). Awareness of potential bias is crucial for maintaining internal validity and supporting

translational research. Replication in different populations and cohorts will strengthen the generalisability of findings and enhance the evidence base in longitudinal research.

4.7. References

Adams, M. J., Hill, W. D., Howard, D. M., Dashti, H. S., Davis, K. A. S., Campbell, A., Clarke, T.-K., Deary, I. J., Hayward, C., Porteous, D., Hotopf, M., & McIntosh, A. M. (2020). Factors associated with sharing e-mail information and mental health survey participation in large population cohorts. *International Journal of Epidemiology*, *49*(2), 410–421. https://doi.org/10.1093/ije/dyz134

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*(3), 317–332. https://doi.org/10.1007/BF02294359

Ask, H., Corfield, E., Biele, G., Tilling, K., McAdams, T. A., Badini, I., Davies, N. M., Davey Smith, G., Richardson, T. G., Cornish, R., Tesli, M., Askeland, R. B., Magnus, P., Andreassen, O. A., Njølstad, P. R., Ystrøm, E., Stoltenberg, C., Reichborn-Kjennerud, T., Hannigan, L. J., & Havdahl, A. (in prep). *The genomic basis of continued participation in the Norwegian Mother, Father and Child Cohort Study*.

Biele, G., Gustavson, K., Czajkowski, N. O., Nilsen, R. M., Reichborn-Kjennerud, T., Magnus, P. M., Stoltenberg, C., & Aase, H. (2019). Bias from self selection and loss to follow-up in prospective cohort studies. *European Journal of Epidemiology*, *34*(10), 927–938. https://doi.org/10.1007/s10654-019-00550-1

Bliddal, M., Liew, Z., Pottegård, A., Kirkegaard, H., Olsen, J., & Nohr, E. A. (2018). Examining Nonparticipation in the Maternal Follow-up Within the Danish National Birth Cohort. *American Journal of Epidemiology*, *187*(7), 1511–1519. https://doi.org/10.1093/aje/kwy002

Cornish, R. P., Macleod, J., Boyd, A., & Tilling, K. (2020). Factors associated with participation over time in the Avon Longitudinal Study of Parents and Children: A study using linked education and primary care data. *International Journal of Epidemiology*, *dyaa192*. https://doi.org/10.1093/ije/dyaa192

Dupuis, M., Strippoli, M.-P. F., Gholam-Rezaee, M., Preisig, M., & Vandeleur, C. L. (2019). Mental disorders, attrition at follow-up, and questionnaire non-completion in epidemiologic research. Illustrations from the CoLaus|PsyCoLaus study. *International*

Journal of Methods in Psychiatric Research, *28*(4), e1805. https://doi.org/10.1002/mpr.1805

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A., Ring, S., Nelson, S. M., & Lawlor, D. A. (2013). Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology*, *42*(1), 97–110. https://doi.org/10.1093/ije/dys066

Goldberg, M., Chastang, J. F., Leclerc, A., Zins, M., Bonenfant, S., Bugel, I., Kaniewski, N., Schmaus, A., Niedhammer, I., Piciotti, M., Chevalier, A., Godard, C., & Imbernon, E. (2001). Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: A prospective study of the French GAZEL cohort and its target population. *American Journal of Epidemiology*, *154*(4), 373–384. https://doi.org/10.1093/aje/154.4.373

Graaf, R. de, Bijl, R. V., Smit, F., Ravelli, A., & Vollebergh, W. A. M. (2000). Psychiatric and Sociodemographic Predictors of Attrition in a Longitudinal Study The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *American Journal of Epidemiology*, *152*(11), 1039–1047. https://doi.org/10.1093/aje/152.11.1039

Hara, M., Sasaki, S., Sobue, T., Yamamoto, S., & Tsugane, S. (2002). Comparison of cause-specific mortality between respondents and nonrespondents in a population-based prospective study: Ten-year follow-up of JPHC Study Cohort I. *Journal of Clinical Epidemiology*, *55*(2), 150–156. https://doi.org/10.1016/S0895-4356(01)00431-0

Hernán, M. A., Hernández-Díaz, S., & Robins, J. M. (2004). A structural approach to selection bias. *Epidemiology (Cambridge, Mass.)*, *15*(5), 615–625. https://doi.org/10.1097/01.ede.0000135174.63482.43

Howe, L. D., Tilling, K., Galobardes, B., & Lawlor, D. A. (2013). Loss to follow-up in cohort studies: Bias in estimates of socioeconomic inequalities. *Epidemiology (Cambridge, Mass.)*, *24*(1), 1–9. https://doi.org/10.1097/EDE.0b013e31827623b1

Lamers, F., Milaneschi, Y., de Jonge, P., Giltay, E. J., & Penninx, B. W. J. H. (2018). Metabolic and inflammatory markers: Associations with individual depressive symptoms. *Psychological Medicine*, *48*(7), 1102–1110. https://doi.org/10.1017/S0033291717002483

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

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature*, *456*(7218), Article 7218. https://doi.org/10.1038/456018a

McAdams, T. A., Rijsdijk, F. V., Zavos, H. M. S., & Pingault, J.-B. (2021). Twins and Causal Inference: Leveraging Nature's Experiment. *Cold Spring Harbor Perspectives in Medicine*, *11*(6), a039552. https://doi.org/10.1101/cshperspect.a039552

Metten, M.-A., Costet, N., Multigner, L., Viel, J.-F., & Chauvet, G. (2022). Inverse probability weighting to handle attrition in cohort studies: Some guidance and a call for caution. *BMC Medical Research Methodology*, *22*(1), 45. https://doi.org/10.1186/s12874-022-01533-9

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. *International Journal of Epidemiology*, *47*(1), 226–235. https://doi.org/10.1093/ije/dyx206

National Institute for Health Research. (2023). *Patient and Public Involvement (PPI)*. Research Design Service South Central. https://www.rds-sc.nihr.ac.uk/ppi-information-resources/

Nilsen, R. M., Vollset, S. E., Gjessing, H. K., Skjaerven, R., Melve, K. K., Schreuder, P., Alsaker, E. R., Haug, K., Daltveit, A. K., & Magnus, P. (2009). Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatric and Perinatal Epidemiology*, *23*(6), 597–608. https://doi.org/10.1111/j.1365-3016.2009.01062.x

Nilsen, T. S., Knudsen, G. P., Gervin, K., Brandt, I., Røysamb, E., Tambs, K., Orstavik, R., Lyle, R., Reichborn-Kjennerud, T., Magnus, P., & Harris, J. R. (2013). The Norwegian Twin Registry from a public health perspective: A research update. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, *16*(1), 285–295. https://doi.org/10.1017/thg.2012.117

Nohr, E. A., Frydenberg, M., Henriksen, T. B., & Olsen, J. (2006). Does low participation in cohort studies induce bias? *Epidemiology (Cambridge, Mass.)*, *17*(4), 413–418. https://doi.org/10.1097/01.ede.0000220549.14177.60

Nohr, E. A., & Liew, Z. (2018). How to investigate and adjust for selection bias in cohort studies. *Acta Obstetricia et Gynecologica Scandinavica*, *97*(4), 407–416. https://doi.org/10.1111/aogs.13319

Pizzi, C., De Stavola, B., Merletti, F., Bellocco, R., dos Santos Silva, I., Pearce, N., & Richiardi, L. (2011). Sample selection and validity of exposure-disease association estimates in cohort studies. *Journal of Epidemiology and Community Health*, *65*(5), 407–411. https://doi.org/10.1136/jech.2009.107185

Relton, C. L., Gaunt, T., McArdle, W., Ho, K., Duggirala, A., Shihab, H., Woodward, G., Lyttleton, O., Evans, D. M., Reik, W., Paul, Y.-L., Ficz, G., Ozanne, S. E., Wipat, A., Flanagan, K., Lister, A., Heijmans, B. T., Ring, S. M., & Davey Smith, G. (2015). Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *International Journal of Epidemiology*, *44*(4), 1181–1190. https://doi.org/10.1093/ije/dyv072

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119–133. https://doi.org/10.1093/bib/3.2.119

Smith, L. H. (2020). Selection Mechanisms and Their Consequences: Understanding and Addressing Selection Bias. *Current Epidemiology Reports*, *7*(4), 179–189. https://doi.org/10.1007/s40471-020-00241-6

Statistics Norway. (2013, January 18). *Norwegian Standard Classification of Education*. Ssb.No. https://www.ssb.no/en/utdanning/norwegian-standard-classification-of-education

Strand, B. H., Dalgard, O. S., Tambs, K., & Rognerud, M. (2009). Measuring the mental health status of the Norwegian population: A comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). *Nordic Journal of Psychiatry*. https://doi.org/10.1080/08039480310000932

Sullivan, G. M., & Feinn, R. (2012). Using Effect Size—Or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Tambs, K., & Moum, T. (1993). How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatrica Scandinavica*, *87*(5), 364–367. https://doi.org/10.1111/j.1600-0447.1993.tb03388.x

Taylor, A. E., Jones, H. J., Sallis, H., Euesden, J., Stergiakouli, E., Davies, N. M., Zammit, S., Lawlor, D. A., Munafò, M. R., Davey Smith, G., & Tilling, K. (2018). Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, *47*(4), 1207–1216. https://doi.org/10.1093/ije/dyy060

Tyrrell, J., Zheng, J., Beaumont, R., Hinton, K., Richardson, T. G., Wood, A. R., Davey Smith, G., Frayling, T. M., & Tilling, K. (2021). Genetic predictors of participation in optional components of UK Biobank. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21073-y

Tyrrell, J., Zheng, J., Beaumont, R., Hinton, K., Richardson, T. G., Wood, A. R., Smith, G. D., Frayling, T. M., & Tilling, K. (2020). Genetic predictors of participation in optional components of UK Biobank. *BioRxiv*, 2020.02.10.941328. https://doi.org/10.1101/2020.02.10.941328

Vejrup, K., Magnus, P., & Magnus, M. (2022). Lost to follow-up in the Norwegian mother, father and child cohort study. *Paediatric and Perinatal Epidemiology*, *36*(2), 300–309. https://doi.org/10.1111/ppe.12821

Winding, T. N., Andersen, J. H., Labriola, M., & Nohr, E. A. (2014). Initial non-participation and loss to follow-up in a Danish youth cohort: Implications for relative risk estimates. *Journal of Epidemiology and Community Health*, *68*(2), 137–144. https://doi.org/10.1136/jech-2013-202707

Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T., & Lamberts, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *The British Journal of Psychiatry*, *195*(3), 249–256. https://doi.org/10.1192/bjp.bp.108.053751

5. Discussion

The aim of this thesis was to explore the nature of the relationship between socioeconomic status (SES) and mental health. Specifically, I sought to examine how socioeconomic factors affect the relative importance of genetic and environmental influences on mental health outcomes in childhood (namely, emotional and behavioural symptoms) and adulthood (namely, depression symptoms). Additionally, I aimed to explore the influence of selection bias in family data and its potential impact on phenotypic and genetic associations between socioeconomic factors and mental health outcomes. I approached these aims in the context of three challenges of gene-environment interaction (GxE) research on socioeconomic inequalities in mental health, discussed below.

5.1. Overview of findings in the context of thesis aims

5.1.1. Improving understanding of the mechanisms driving socioeconomic disparities in children's mental health: gene-environment interaction in the presence of gene-environment correlation

In chapter 2 of this thesis, I aimed to contribute new knowledge on the role that parents' socioeconomic factors play in shaping the origins of child emotional and behavioural problems. I applied a novel genetically informed design using extended family data from the Norwegian, Mother, Father and Child Cohort Study (MoBa) (Magnus et al., 2016) to investigate whether parental socioeconomic factors moderate the genetic and environmental influences on child emotional and behavioural symptoms (GxE). In contrast to previous twin studies in children (Burt et al., 2016, 2020; Hendriks et al., 2019; Middeldorp et al., 2014; Tuvblad et al., 2006), I was able to test for gene-environment interaction (GxE) *in the presence of modelled gene-environment correlation* (rGE) for environmental moderators that are necessarily shared between children growing up in the same family.

My results showed that family socioeconomic factors moderate the influence of familial and non-shared environmental influences on child emotional and behavioural symptoms. Results were consistent with prior evidence indicating that shared environmental influences on child emotional problems have a greater influence among children from families with lower SES

(consistent with the bioecological framework; Bronfenbrenner & Ceci, 1994). Interestingly, my ability to model the covariance between socioeconomic factors and child outcomes revealed that shared genetic factors explained more of the variance in child behavioural symptoms for families at lower levels of socioeconomic status (consistent with the diathesis stress framework; Monroe & Simons, 1991). This increased heritability was driven by genetic variance shared between maternal SES indicators and child behavioural symptoms that would have been regressed out of previous GxE analyses. Thus, it is possible that previous studies have not detected increased genetic influences at lower ends of the socioeconomic distribution because the genetic effects are shared between parent SES and child psychopathology. Findings from this study also indicated that the presence and pattern of moderation can vary depending on the measure used to index family socioeconomic conditions. Analysing the moderating effects of individual socioeconomic indicators, separately for mothers' and fathers', revealed that maternal and paternal socioeconomic factors may partly reflect individual characteristics of the parents. This suggests that parents' socioeconomic factors may influence child behaviour via different pathways, highlighting the multidimensional nature of socioeconomic status.

5.1.2. Improving understanding of the mechanisms underpinning socioeconomic disparities in mental health in low- and middle-income populations

In chapter 3, I aimed to contribute new knowledge on the mechanisms by which socioeconomic conditions influence depression in low- and middle-income countries. To my knowledge, this is the first study to use biometric moderation modelling to investigate whether socioeconomic factors moderate aetiological influences on depression symptoms in a Sri Lankan population.

Findings revealed that lower standard of living, poor educational attainment, and financial strain were associated with higher levels of depression symptoms, consistent with previous research. The study also identified sex differences in the genetic influences on standard of living, with a small genetic contribution observed in females but not in males. Genetic influences on standard of living in females may reflect lower variation in environmental exposures due do cultural gender limitations. For example, the majority of working age females in Sri Lanka are not in salary-based employment (~73%) with the main reasons

reported being a caregiver, family work, and household activities (Department of Census and Statistics Sri Lanka, 2019). Shared environmental factors environmental factors accounted for the majority of variance in standard of living in both males and females, potentially reflecting socio-cultural norms (e.g., higher prevalence of extended, multigenerational, family households; Maselko, 2017) and/or less equal access or opportunity in employment sectors in Sri Lanka. Educational attainment showed moderate heritability in both sexes. This replicates findings in other countries and is often interpreted as an index of meritocracy – whereby educational attainment is influenced by an individual's genetic propensity to perform well in educational settings (Rimfeld et al., 2018). We found evidence that aetiological influences unique to depression symptoms were moderated by standard of living and educational attainment. However, it was not possible to determine whether moderation operated through genetic and/or environmental mechanisms. Nevertheless, the moderation of depression by standard of living and educational attainment is consistent with the notion that social causation processes play a role in the observed association between socioeconomic status and depression symptoms across different populations (Dohrenwend et al., 1992; Lund & Cois, 2018; South & Krueger, 2011). Findings also showed that the shared variance components between socioeconomic factors and depression symptoms were zero across the entire socioeconomic distribution, suggesting that socioeconomic factors may have a main effect on the aetiology of depression independently of shared aetiological influences.

The study's strengths include a large representative population-based twin and singleton sample in the Colombo District in Sri Lanka and the use of standard measures of asset and education. Limitations of the study include potential underreporting of socioeconomic factors and depression symptoms due to sensitivity, privacy, and stigma (Lorant et al., 2007), lack of investigation into age-related patterns of moderating effects by socioeconomic conditions, potential regional representativeness concerns, and the reliance on assumptions in the twin method that may impact result validity (Rijsdijk & Sham, 2002). Overall, findings contribute to our understanding of the complex relationship between socioeconomic factors and depression, with implications for future cross-cultural research and interventions addressing social disparities in depression.

5.1.3. Considering selection bias in family data

In my final study, I investigated the impact of selective attrition on phenotypic and genetic exposure-outcome associations in the MoBa cohort. Findings showed that participation rates declined over time, with higher rates among older mothers, first-time mothers, and those with higher educational attainment. Conversely, attrition was linked to higher levels of internalising symptoms among mothers. In line with previous findings (Howe et al., 2013), selective participation had an effect on the association between maternal educational attainment and internalising symptoms, with weaker associations observed among those with higher participation rates. This suggests the possibility of collider bias, indicating that conditioning on continued participation, a potential collider, may have introduced bias into the estimate of the association between educational attainment and internalising symptoms. However, no moderation effects were found for the association between maternal income rank and internalising symptoms or for the associations between maternal socioeconomic factors and offspring outcomes, indicating that the impact of selective participation on these specific associations was minimal. Genetic factors played a significant role in continued participation, and genetic correlations were observed between educational attainment or internalising symptoms and continued participation. Consistent with our study findings, previous research using genotype data has identified genetic variants associated with ongoing study participation, and genetic overlap between participation and educational attainment and depression (Adams et al., 2020; Ask et al., 2021; Taylor et al., 2018; Tyrrell et al., 2020). The impact of selection bias on the aetiological relationship between maternal educational attainment and internalising symptoms in MoBa was minimal, as the observed genetic correlation remained consistent across the full sample and a subset with follow-up data.

Limitations of the study include a specific focus on selection bias at follow-up rather than at baseline, the need for future research to link registry data for the entire target population to investigate biases, and the potential lack of generalisability to studies with different selection criteria, recruitment methods, or cultural/contextual factors. Additionally, the study did not explore all possible traits associated with participation, and future studies should examine the impact of selection bias on associations between other socioeconomic factors and mental health outcomes. Overall, the findings from this study highlights the factors associated with

continued participation in the MoBa cohort and the potential bias introduced by selective attrition on phenotypic exposure-outcome associations. The findings underscore the importance of identifying predictors of participation and considering potential bias to maintain internal validity and support translational research. Replication of these findings in diverse populations and cohorts will enhance the generalisability of the results and contribute to the evidence base in longitudinal research. Such research efforts will help to identify under-represented groups where community engagement and involvement initiatives could be focussed.

5.2. General limitations

I have discussed the limitations associated with each research study throughout the chapters of this thesis. In the following paragraphs, I consider some of the key limitations of this research in sum. These limitations reflect challenges that apply to all studies exploring the role of gene-environment interplay in the relationship between socioeconomic status and mental health, including my presented research.

5.2.1. Generalisability

The findings from chapter 4 in this thesis showed that the MoBa cohort (used in chapters 2 and 4) is characterised by self-selection, with ongoing study participants more likely to be older, first-born mothers, have higher educational attainment, and lower self-reported internalising symptoms, compared to those lost to follow-up and to the national average (Magnus et al., 2016). Similarly, in the CoTASS sample, which represents the Colombo District, selective participation based on factors such as age, gender, ethnicity and twin status is evident (Jayaweera et al., 2018). Additionally, mental health data used across chapters were positively skewed, indicating low levels of symptoms among study participants. Such selection biases are a pervasive issue for all scientific research that relies upon voluntary participation. Selection biases can be influenced by factors operating at a global (e.g., inequality in research funding and infrastructure may lead to the clustering of cohorts in developed countries), national (e.g., cultural and societal factors may lead to gender and racial biases in study participation), institutional (e.g., lack of diversity among scientists who design and implement research protocol may lead to biases in research) and individual (e.g., greater exposure to life stressors may lead to lower capacity to engage in research) level.

My results suggest that selection bias may only influence phenotypic associations in cases where both the predictor and outcome predict study participation. Collider bias occurs when conditioning on a variable, in this case, continued participation, that is influenced by both the predictor (educational attainment) and the outcome (internalising symptoms). The weaker association observed among those with above-average and average participation rates in chapter 4 suggests that conditioning on continued participation, a potential collider, may have introduced bias into the estimate of the phenotypic association between participants' educational attainment and internalising symptoms in MoBa. Such bias could be problematic particularly if similar characteristics related to participation in different studies as the results might be reproducible but may equally be biased. Given that internalising symptoms do not predict follow-up study participation in CoTASS, this could suggest a minimal impact of selection bias on association estimates between socioeconomic factors and depression presented in chapter 3. Further, my results from chapter 4 also showed that attrition may have a limited impact on estimates of genetic and environmental sources of variance and covariance between socioeconomic factors and internalising problem. This provides additional reassurance to the validity of results presented throughout the chapters in this thesis.

The restricted range of participants in terms of socioeconomic status and mental health raises concerns regarding the generalisability of gene-environment interaction (GxE) findings. Without a representative sample covering the full distribution of these variables, it is difficult to fully understand patterns of GxE at the extreme ends of the distribution (Purcell, 2002). Furthermore, the lack of representativeness in samples means that the voices of marginalised individuals are not being heard or adequately represented. While there are statistical corrections available to address biases and make inferences on the entire target population, such as weighting methods or imputation techniques (Metten et al., 2022), these approaches do not address broader issues of representativeness.

As in all research, the findings presented in this thesis are specific to the context in which they were derived, and caution should always be exercised when generalising results to different contexts. For instance, Norway, a country with a strong welfare state, may exhibit different

associations between socioeconomic status and mental health problems compared to Sri Lanka, which may have a weaker safety net or variation in other potentially important socioeconomic aspects (e.g., lower labour force participation rates, especially among females [Department of Census and Statistics Sri Lank, 2022]). Importantly, population estimates from twin and family-based studies are context- and time-specific, and do not represent “what could be” but only “what is" at one point in time, in a particular setting. Differences in social, cultural and economic contexts within and between different regions can shape the relationships between variables, highlighting the importance of considering the broader context when interpreting and generalising study results.

The lack of diversity among the samples used in this thesis and in existing mental health research, as well as among research communities, perpetuates inequality in mental health and in human science research (Cooper et al., 2013; Maura & Weisman de Mamani, 2017). Studies from LMIC are severely under-represented in behavioural genetic research, a problem highlighted in a review of all twin studies conducted over the past 50 years (Polderman et al., 2015). The continents of Africa, South America and Asia represented only 0.2-5% of published twin research. Similarly, the vast majority of genome-wide association studies (~88%) have been conducted on individuals of European ancestry, with African, African-American or Afro-Caribbean, Hispanic and Latin American, and Asian ancestry groups represented in only 1.0-2.8% of published studies (Mills & Rahal, 2020). Researchers must acknowledge these limitations and actively promote inclusivity by conducting diverse research, developing culturally appropriate mental health measures, improving participant recruitment strategies, and seeking global funding opportunities. Failure to do so risks perpetuating a cycle of biased scientific evidence, and subsequently biased policy and intervention, that caters more towards privileged groups.

Researchers can adopt strategies to improve study participation and strive for better representation. These strategies may include targeted recruitment efforts to reach underrepresented populations, creating inclusive study environments, addressing barriers to participation, and incorporating diverse perspectives in study design and interpretation of findings. Patient and Public Involvement and Engagement (PPIE), also known as Community Engagement and Involvement (CEI), has the potential to assist researchers, healthcare

professionals, and stakeholders in effectively planning and conducting research studies, as well as implementing health policies, that better align with the specific requirements of the community (National Institute for Health Research, 2023). It is crucial for researchers to acknowledge and actively address the limitations of sample representativeness and work towards improving inclusivity and diversity of research studies. By doing so, researchers can ensure that findings are more applicable and relevant to a broader range of individuals, and that voices of marginalised population groups are heard and represented.

5.2.2. Measurement

5.2.2.1. Reporter biases

Across the chapters in this thesis, data on emotional and behavioural symptoms in children, and internalising symptoms in adults, were collected using parent-report or self-report questionnaire items. Compared to other data collection methods, questionnaires are relatively inexpensive and simple to administer across the large samples that are required for genetically informed research. In contrast to other sources of information (e.g., national administrative data), however, questionnaires rely solely on participant perspectives and ability and/or willingness to respond accurately. Self-report data are subject to a range of unwanted biases that can arise from inaccurate or biased responding by participants.

Response biases such as social desirability bias have been shown to affect the validity of self-report measures of socioeconomic factors and psychiatric symptoms due to the sensitive and/or private nature and stigma associated with reporting them. For example, there is evidence of social desirability bias in reporting personal income; individuals at the lower end of the wage distribution are more likely to overreport their income, whereas the highest earners are more likely to underreport their income (Angel et al., 2019). In chapters 2 and 4, the use of register data to index socioeconomic conditions is an important strength because it means that social desirability bias does not influence the socioeconomic measures used in the (MoBa) analyses. Additionally, the use of register data allowed me to retain individuals in the analyses who did not provide self-report income and/or education MoBa data. However, the availability and utility of such data is often limited or non-existent particularly from low-

and middle-income populations (LMIC). Nevertheless, the wide availability of asset index data and educational attainment in many studies and comparable data across countries is an important strength of the study in chapter 3 because it facilitates comparative research. Research has shown that in many LMIC, such as Sri Lanka, mental health difficulties can carry additional cultural stigma and individuals may be more inclined to mask psychiatric rather than, for example, physical health conditions (Mascayano et al., 2015; Samarasekara et al., 2012). Thus, socio-cultural stigmas associated with socioeconomic conditions and mental health may have hindered the reporting of them in chapter 3. While this could result in biased prevalence estimates (Munafò et al., 2018), it is unlikely that results from these analyses were biased by selection, given that neither socioeconomic status (indicated by educational attainment or standard of living) or depression predicted follow-up participation in CoTASS-2 (Jayaweera et al., 2018).

5.2.2.2. Measurement tools and procedure

Another potential limitation is that the questionnaire data used in this thesis provide only crude measures of participants' mental health symptoms. For example, in chapter 3 the five-item version of the SCL-5 was used to measure mothers' internalising symptoms and has been shown to have relatively low reliability. Nevertheless, prior evidence suggests overall good psychometric properties of this scale (Schmalbach et al., 2021). The socioeconomic indicators used throughout the chapters of this thesis only capture certain aspects of participants' socioeconomic conditions at the time of assessment. Although different socioeconomic indicators are correlated, each has been shown to measure a distinctive aspect of the environment and may reflect different pathways through which mental health is affected (Duncan & Magnuson, 2003; Festin et al., 2017; Howe et al., 2012). Moreover, the chapters in this thesis focussed specifically on individual-level socioeconomic indicators, rather than neighbourhood- or country-level indicators of socioeconomic status. As a result, the understanding of how broader structural socioeconomic inequalities can impact the development of internalising and externalising problems in children and adults may be limited.

5.2.3. Dependence of gene-environment interactions on scaling

In practice, the ability to test theoretically implied GxE interactions is constrained by the dependence of interaction tests on the observed distributions of the traits (Murray et al., 2016; Van Hulle & Rathouz, 2015). The presence of non-normality in an observed phenotypic distribution can lead to incorrect inferences regarding the presence of GxE. Non-normality could arise for reasons such as failing to sample individuals with the lowest or highest trait levels from the population or if the measurement instrument used does not capture the full range of variation in the trait. Many traits of interest in behavioural genetic research, as well as psychological research in general, show skewed score distributions. Typical behavioural genetic approaches to correcting for non-normality include non-linear transformations of sum scores or estimating factor scores from an item response theory (IRT) model (Murray et al., 2016). Transformation of non-normal sum-scores to normality has been shown to improve false-positive rates and provide less biased parameter estimates in GxE tests. Data on mental health symptoms used across the empirical chapters of this thesis were positively skewed and subsequently corrected using log-linear transformations.

GxE estimates depend on the degrees of individual differences in a trait at different levels of the moderator, and thus, failing to adequately sample the full socioeconomic distribution from the population can reduce statistical power to detect GxE effects and lead to incorrect inferences of the mechanism of interaction. For example, the MoBa dataset used in chapters 2 and 4 is characterised by self-selection and participants have been found to have higher levels of SES compared to those who did not participate. Less variation at lower levels of SES can alter the shape and significance of interactions at this end of the distribution. Thus, it is important for future work to consider the distributions of SES indicators when testing for theoretically implied GxE interactions, as well as applying appropriate transformations to skewed trait scores. There are ongoing efforts to link MoBa questionnaire data with information from a range of administrative registers, including demographic, medical, income and wealth, education, and criminal records. These linkages provide further opportunities to assess selection characteristics of the sample and implement methods to correct for bias.

5.2.4. Cross-sectional data

The empirical studies presented in this thesis adopt a cross-sectional design, limiting the findings to the concurrent aetiology and relationships between variables, specific to the time and/or developmental period at which they were assessed. The nature and magnitude of associations between socioeconomic indicators and mental health outcomes is influenced by various contextual factors and may exhibit age-related changes. For example, studies have identified time-varying associations between family socioeconomic conditions and trajectories of emotional and behavioural problems. Developmental differences in moderation could be expected given that the aetiology and trajectories of mental health problems change across development (Hannigan et al., 2017). This is especially relevant for chapter 2, which is focussed on childhood mental health outcomes, given that various genetic and environmental factors may become more or less influential as certain behaviours and cognitive processes become involved in driving the symptoms of emotional and behavioural problems at different stages of development.

Socioeconomic status is not a static construct, and research suggests that mental health and other adverse outcomes may be worsened by sharp income fluctuations caused by economic insecurity (Hardy & Ziliak, 2014; Jacobs & Hacker, 2008; Miller et al., 2021; Whitfield et al., 2021). Longer-term financial instability may also influence the aetiology of mental health problems in ways which I have not explored in the chapters of this thesis. Longitudinal studies have shown links between increased income *volatility* and poorer socioeconomic and health outcomes (Miller et al., 2021; Morrissey et al., 2020; Whitfield et al., 2021). In recent years, household income has become increasingly unstable, especially in families with children and among low-wage workers (Hardy & Ziliak, 2014; Hill et al., 2013). COVID-19 and other economic crises, including the recent economic crisis in Sri Lanka, have further exacerbated inequalities in social, economic and mental health outcomes (Kola et al., 2021; Matthias & Jayasinghe, 2022; Santomauro et al., 2021). This has implications for the interpretation of the findings presented throughout this thesis because estimates of genetic and environmental sources of individual differences and underlying causal mechanisms can change over time and as circumstances change.

5.2.5. Sample size

An important issue for GxE research in general is that many GxE twin studies are underpowered due to sample size, which raises concerns as to the likelihood that detected interaction effects are true positives. Ideally, adequate power to detect true interaction effects in favour of alternative models requires sample sizes of at least 2,000 twin pairs (Van Hulle & Rathouz, 2015). A noteworthy strength of the study conducted in chapter 2 is the use of a large prospective cohort study (MoBa), which included over 16,000 extended mother-child and father-child extended families. Although the CoTASS dataset used in chapter 3 included a larger sample size than possible comparison published studies, my results suggest that this research would have benefited from larger samples to refine findings. Confidence intervals around the moderated parameters were large and non-significant. However, results from post-hoc model fitting analyses suggest moderation on the aetiological influences unique to depression symptoms (i.e., moderation on the variance components not shared with the socioeconomic indicator) and are consistent with the moderation effects observed in the interaction graphs (Chapter 3, figure 2).

5.3. Avenues for future research

5.3.1. Exploring the relationship between socioeconomic status and mental health across time and populations

An interesting future direction could be to explore the nature of the relationship between socioeconomic conditions and mental health outcomes across time. This could involve examining whether longitudinal changes in family socioeconomic conditions are associated with developmental change in offspring emotional and behavioural difficulties across childhood and adolescence. In chapter 2, I found novel evidence of moderation of genetic variance that is shared between maternal income and offspring behavioural symptoms in early childhood. However, it is possible for developmental differences in moderation given that the aetiology and trajectories of mental health problems change across development. For example, emotional and behavioural problems tend to increase with the onset of puberty and transition into adolescence (Costello et al., 2005; Scaramella et al., 1999). At the same time, adolescents become increasingly aware of their own social standing (Harold & Conger, 1997). Thus, examination of these later developmental periods may reveal different processes

of gene-environment interplay underlying observed relations between socioeconomic conditions and mental health. Work to combine longitudinal models with an extended family design has already begun, in attempts to explore longitudinal co-development of mothers' internalising and offspring temperament characteristics during early development (Ahmadzadeh et al., 2023). Future studies could next consider using similar large-scale genetically informative data to build on these models to explore how the dynamics of different measures of the family socioeconomic environment relate to *trajectories* of aetiologic moderation on emotional and behavioural problems across child developmental periods. In addition to exploring changes in GxE across development, future research employing a longitudinal design could also explore stability in GxE effects that persist across time.

Another future direction could be to examine the direction of the relationship between socioeconomic status and mental health outcomes. For example, my findings from chapter 3 provide support for both social selection (significant genetic influences on SES indicators) and preliminary support for social causation (GxE). A next step could be to expand my analyses to include data from the first wave (CoTASS-1) and, when available, data collected at the third assessment (CoTASS-3). This would allow for investigations of the direction of effects between socioeconomic status and mental health. For example, longitudinal phenotypic research has found evidence for a reciprocal relationship between SES and depression, suggesting social causation and social selection act simultaneously to reinforce cycles of socioeconomic disadvantage and depression (Jin et al., 2020; Lund & Cois, 2018). Using longitudinal genetically informed data, biometric autoregressive cross-lagged models could be used to examine whether socioeconomic circumstances prospectively predict mental health difficulties, or vice versa (i.e., mental illness predicts socioeconomic status), after accounting for shared aetiology. Versions of this model have been used to demonstrate, for example, that social anxiety disorder prospectively predicts alcohol use disorder, and not vice versa, after correction for shared aetiologies (Torvik et al., 2019). A potential is that the socioeconomic status of participants may not change enough over time to be able to conduct such analyses. For example, individuals in CoTASS showed similar rates of full-time employment at the second-wave of assessment (~56% in CoTASS-2) as reported at the first-wave of assessment (~53% in CoTASS-1) (Jayaweera et al., 2018; Siribaddana et al., 2008).

Following from this, a broad future avenue is the inclusion of more diverse samples within behaviour genetics and science in general. The limited representation of ancestral, geographic, and demographic diversity in the existing data poses a significant obstacle to genetic research. Underrepresentation of diverse populations limits our understanding of the mechanisms through which socioeconomic and other social risk factors impact mental health and may exacerbate existing health inequalities. As discussed in chapter 4, those who are less advantaged and less healthy are often under-represented in studies, leading to samples that are generally 'healthier and wealthier' than the intended study population (Howe et al., 2013; Jacobsen et al., 2010; Nohr & Liew, 2018). My findings indicate that selective participation at follow-up assessments (i.e., attrition) may result in underestimation of socioeconomic inequalities in internalising symptoms, and similar biases in several health-related outcomes have been previously reported (Howe et al., 2013). Future studies may consider applying methods to adjust for results in the study sample to match results in the target population, such as inverse probability weighting. Ultimately, however, collection of more representative and diverse data would be a preferable solution. Another serious issue is that genetically informative research has been largely restricted to high-income populations (Polderman et al., 2015). Further, bias towards Western populations is also problematic because estimates of genetic and environmental sources of individual differences are specific to a population at a particular time period and have been shown to vary cross-culturally (Hur, 2008; Zavos et al., 2020). In this thesis, I used a distinct Sri Lankan sample mentioned in Chapter 3, which stands as a singular twin and singleton registry within a South Asian context. However, with most twin studies still aggregating in Western populations, there is an ongoing global necessity for more twin registries. Such research efforts would help to provide new insights into the genetic and environmental factors, as well as their interplay, underlying complex traits.

5.3.2. Modelling broader family systems

A next step could be to expand my analyses to model broader family systems. For example, this could involve examining associations between indicators of family socioeconomic conditions and offspring mental health in models that also account for associations between two parents. In chapter 2, I modelled extended family units separately for sibling pairs of mothers or fathers and their children. In doing so, I was able to disentangle the effects of

mother and father socioeconomic factors on the aetiologies of child emotional and behavioural problems. Results indicated that aetiological influences on child outcomes were differentially moderated by maternal versus paternal socioeconomic indicators, which could suggest that these indicators may partly be reflective of individual traits of the parents and may affect child outcomes via different pathways. While I modelled mothers and fathers separately, future analyses could also seek to model them simultaneously as an overall measure of family socioeconomic status and investigate whether moderation effects differ between individual and composite measures of SES.

Future studies could also consider modelling both parents per child to examine partner associations between socioeconomic indicators and potential effects of assortative mating on parent-offspring correlations. This may be especially relevant when examining the effects of parents' educational attainment on child mental health outcomes as partners have been found to strongly resemble each other on educational attainment in previous studies (Greenwood et al., 2014; Torvik et al., 2020, 2022), and in my analyses in chapter 2 (r = 0.45). Torvik et al. (2020) present an extended MCoTS model in which two parents per child are included in a family unit. Specifically, the authors sought to estimate the genetic overlap between partners for educational attainment by comparing phenotypic correlations between parents and their siblings, parents and their partners, and parents and their sibling's partners (i.e., in-laws). This estimate of partner genetic similarity was used to adjust offspring sibling correlations to account for the inflation of sibling genetic overlap as a function of assortative mating among parents. Using this approach, it is possible to model the effects of maternal and paternal educational attainment separately on the offspring outcome after accounting for intergenerationally shared genetic and environmental influences. Future studies could seek to extend this model to include moderation parameters on these intergenerational pathways to explore differences in results between parents.

Linkage to population and health registers may also provide opportunities for future studies to model broader family systems, such as relatives of study participants who did not participate. For example, in chapter 4 I used ancestry information from the Norwegian birth register linked with data from MoBa to identify additional extended families linked by mothers who were cousins. Further, linkage to individual-level register data can be used to identify all

members of a population who were eligible to participate in a study. Work to identify non-participants who were eligible to take part in MoBa has already commenced, in efforts to examine selection biases attributable to differences between participants and non-participants, rather than between continuing participants and drop-outs. Linkage to data from birth records enables identification of pedigree structures within the target population, and this information can be used to conduct genetically informed analyses. This approach could be taken in future research examining the potential impact of bias due to selective initial participation and subsequent attrition on estimated phenotypic and genetic exposure-outcome associations. This may be especially relevant to research on social inequalities in mental health, given that these factors are strongly associated with participation.

5.3.3. Modelling broader social systems

In the chapters of this thesis, I use data on income or assets and educational attainment to index individuals' socioeconomic circumstances. Thus, my analyses focus on the effects of individual or familial socioeconomic disadvantage on mental health. The effects of these socioeconomic factors on health inequalities do not operate in a vacuum, but rather interact with factors operating across multiple levels of the socioeconomic context, such as neighbourhoods, communities, and wider society, institutions, and culture. For example, the deprivation amplification model proposes that socioeconomic factors operating at the individual- and neighbourhood-level may act together to influence risk for mental health problems, such that individuals with low socioeconomic status living in deprived neighbourhoods may have elevated risk compared with individuals of low SES living in less disadvantaged areas (Macintyre et al., 1993; Mann et al., 2022; Maselko, 2017; Stafford & Marmot, 2003; Visser et al., 2021). In addition to the effects of socioeconomic factors, important inequalities also exist across other population demographic groups, such as sex, gender, race and ethnicity (Assari, 2017; United Nations, 2020). The term 'intersectionality' is used to describe the combined effects of multiple social categories on outcomes of interest (Crenshaw, 1991). For example, findings from chapter 3 revealed sex differences in the aetiology of standard of living, with a small contribution of genetic influences in females but not in males. One interpretation of this finding is that genetic influences on standard of living in females may reflect lower variation in environmental exposures due to cultural gender

inequalities in employment. This suggests the need for future research to consider the intersections between socioeconomic and other social categories. Thus, a next step could be to expand my analyses to model broader social systems which consider the effects of multiple social factors.

One approach to do so could be multilevel analysis, which takes into account the natural clustering of one unit of analysis (e.g., individual) within another (e.g., area of residence), and can be used to distinguish between contextual (macro) and compositional (micro) influences. A recent study in MoBa using genotype data linked with national administrative data on school and residential environments used multilevel modelling to test for gene-environment interactions between polygenic indices of educational attainment (EA-PGI) and environmental levels (Cheesman et al., 2022). Findings indicated an interaction between EA-PGI and schools, but not neighbourhood, district or municipality. This approach could be considered in future studies using similar large-scale genetically informative data to explore the causal influences of multiple socioeconomic and social factors on mental health outcomes.

GxE research the potential to improve our understanding of behavioural traits linked to social inequality and aid in developing interventions that can be applied at multiple levels. By identifying genetic and environmental moderation in specific contexts, population-level studies can assess the impact of social and economic policies on individual differences in mental health. For instance, the social control/push framework emphasises the role of societal norms and constraints in shaping behaviour, suggesting policies that strengthen social institutions, support positive family dynamics, and promote community engagement. Investments in education, community centres, and mentorship programs can contribute to a protective social environment. Policies aligned with the diathesis-stress model could involve early interventions, stress reduction initiatives, and community resources as buffers against negative environmental influences. The bioecological framework suggests a need for policies addressing multiple environmental levels, including family, school, and community interventions fostering supportive environments. Differential susceptibility framework considerations may lead to personalised education plans and mental health support. Highlighting social and economic inequalities that contribute to mental health disparities at the population level is crucial. All GxE models emphasise the influential role of social and

environmental contexts in the causes of physical and mental health disorders, reinforcing the importance of upstream societal interventions to improve living and working conditions, such as labour market regulation, protective employment policies and investment in social protection (Macintyre et al., 2018). Effective policies should recognize the dynamic interplay between genetic predispositions and environmental factors, advocating for a comprehensive, flexible, and multifaceted approach that integrates genetic research with social and environmental interventions.

5.3.4. Incorporating genomic data

In the chapters of this thesis, I use twin/family data to infer the role of genetic factors and their interaction with the socioeconomic context. However, recent methodological advances combining family-based methods with genomic data are giving rise to new research opportunities for elucidating the mechanisms through which socioeconomic factors impact mental health. While genomic methods allow for directly testing gene-environment interplay, they cannot yet replace the methods used in this thesis given that they are currently limited to additive effects of common genetic variants tagged on DNA arrays, rather than approximating influence of the entire genome as in twin and family-based methods (Cheesman et al., 2017; Maher, 2008; Plomin, 2022). Moreover, the biometric moderation models used in the chapters of this thesis test for variation not only in genetic influences but in non-genetic influences too. For example, chapter 2 showed that shared environmental influences on child emotional problems have a greater influence among children from lower socioeconomic backgrounds, which would not be detected using genomic-based designs. Identifying contexts in which environmental factors have notable influence on aetiology has the potential to inform intervention strategies to reduce social inequalities in mental health. Nevertheless, there are ways that future studies may consider applying genomic methods to test similar research questions and hypotheses to those explored in this thesis. A particularly exciting prospect of these methods is that they can be integrated with twin and family-based datasets. In recent years, genomic data have been collected on participating members of large-scale family datasets. For example, previous studies have used parent-child genomic data to explore environmentally mediated genetic effects (termed genetic nurture) of parents on child depression symptoms (Cheesman et al., 2020). Further, parent-child genomic data

can be used to control for gene-environment correlation in GxE analyses. In the study by Cheesman et al. (2022) discussed in the previous section, the authors control for gene-environment correlation by regressing out the effects of parental genotype on offspring genotype prior to testing for interaction. Future studies could consider using similar approaches to explore how family socioeconomic conditions moderate genetic risk for psychopathology.

Genome-wide methods have been proposed to identify variance-controlling genes that confer differential sensitivity to the environment (Conley et al., 2018; Paré et al., 2010; Visscher & Posthuma, 2010; Yang et al., 2011). Genetic effects on trait variance can be inferred from variance quantitative trait locus (vQTL) analysis among groups of unrelated or related individuals. Polygenic scores can be calculated based on variance effects found in vQTL analysis to predict environmental sensitivity and can be examined as moderators of specific environmental exposures on outcomes. For example, one study showed that polygenic scores for environmental sensitivity significantly moderated the effects of parenting behaviour on emotional problems (Keers et al., 2016). In individuals with high polygenic environmental sensitivity, negative parenting was associated with increased emotional symptoms, while positive parenting was associated with decreased emotional problems. An interesting future direction could be to investigate whether polygenic scores for environmental sensitivity moderate the effects of socioeconomic factors on mental health outcomes. Work to conduct genome-wide vQTL analysis in MoBa has already commenced, in efforts to calculate polygenic indices for environmental sensitivity based on the whole-genome findings. These polygenic indices can be tested as both a predictor and moderator of child/adult mental health outcomes in future research.

Another future direction could be to expand my analyses in chapter 4 to incorporate genomic approaches. For example, I could use genomic methods to estimate the SNP heritability of attrition (as implemented in the genome-wide complex trait analysis (GCTA) software) (Yang et al., 2011) and examine what proportion of pedigree heritability is accounted for by SNP-based heritability. Future studies may also consider using parent-child genomic data to explore the role of parent(s) versus child in determining continued participation in the study (e.g., trio-GCTA analysis; Eilertsen et al., 2021).

5.4. Final conclusion

In this thesis, I aimed to explore the relationship between socioeconomic status (SES) and mental health, focusing on the role of genetic and environmental influences and their interplay. Through a series of empirical studies, I investigated the impact of socioeconomic factors on mental health outcomes in both childhood and adulthood, considering gene-environment interaction (GxE) and gene-environment correlation (rGE). Additionally, I examined the influence of selection bias in family data and its potential impact on associations between socioeconomic factors and mental health outcomes. The findings of this thesis contribute to a deeper understanding of the mechanisms underlying socioeconomic disparities in mental health. The findings underscore the importance of considering gene-environment interplay and selection biases in understanding socioeconomic disparities in mental health outcomes. Future research should aim to address the identified limitations and explore these relationships in diverse populations and representative samples.

5.5. References

Ahmadzadeh, Y. I., Eilertsen, E. M., Cheesman, R., Rayner, C., Ystrom, E., Hannigan, L. J., & McAdams, T. A. (2023). Mothers' symptoms of anxiety and depression and the development of child temperament: A genetically informative, longitudinal investigation. *JCPP Advances*, *n/a*(n/a), e12171. https://doi.org/10.1002/jcv2.12171

Angel, S., Disslbacher, F., Humer, S., & Schnetzer, M. (2019). What did you really earn last year?: Explaining measurement error in survey income data. *Journal of the Royal Statistical Society Series A*, *182*(4), 1411–1437.

Assari, S. (2017). Social Determinants of Depression: The Intersections of Race, Gender, and Socioeconomic Status. *Brain Sciences*, *7*(12), 156. https://doi.org/10.3390/brainsci7120156

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, *101*(4), 568–586. https://doi.org/10.1037/0033-295x.101.4.568

Burt, S. A., Klump, K. L., Gorman-Smith, D., & Neiderhiser, J. M. (2016). Neighborhood Disadvantage Alters the Origins of Children's Nonaggressive Conduct Problems.

Clinical Psychological Science : A Journal of the Association for Psychological Science, *4*(3), 511–526. https://doi.org/10.1177/2167702615618164

Burt, S. A., Pearson, A. L., Carroll, S., Klump, K. L., & Neiderhiser, J. M. (2020). Child Antisocial Behavior Is more Environmental in Origin in Disadvantaged Neighborhoods: Evidence Across Residents' Perceptions and Geographic Scales in Two Samples. *Journal of Abnormal Child Psychology*, *48*(2), 265–276. https://doi.org/10.1007/s10802-019-00587-6

Cheesman, R., Borgen, N. T., Lyngstad, T. H., Eilertsen, E. M., Ayorech, Z., Torvik, F. A., Andreassen, O. A., Zachrisson, H. D., & Ystrom, E. (2022). A population-wide gene-environment interaction study on how genes, schools, and residential areas shape achievement. *Npj Science of Learning*, *7*(1), Article 1. https://doi.org/10.1038/s41539-022-00145-8

Cheesman, R., Eilertsen, E. M., Ahmadzadeh, Y. I., Gjerde, L. C., Hannigan, L. J., Havdahl, A., Young, A. I., Eley, T. C., Njølstad, P. R., Magnus, P., Andreassen, O. A., Ystrom, E., & McAdams, T. A. (2020). How important are parents in the development of child anxiety and depression? A genomic analysis of parent-offspring trios in the Norwegian Mother Father and Child Cohort Study (MoBa). *BMC Medicine*, *18*(1), 284. https://doi.org/10.1186/s12916-020-01760-1

Cheesman, R., Selzam, S., Ronald, A., Dale, P. S., McAdams, T. A., Eley, T. C., & Plomin, R. (2017). Childhood behaviour problems show the greatest gap between DNA-based and twin heritability. *Translational Psychiatry*, *7*(12), Article 12. https://doi.org/10.1038/s41398-017-0046-x

Conley, D., Johnson, R., Domingue, B., Dawes, C., Boardman, J., & Siegal, M. (2018). A sibling method for identifying vQTLs. *PLOS ONE*, *13*(4), e0194541. https://doi.org/10.1371/journal.pone.0194541

Cooper, C., Spiers, N., Livingston, G., Jenkins, R., Meltzer, H., Brugha, T., McManus, S., Weich, S., & Bebbington, P. (2013). Ethnic inequalities in the use of health services for common mental disorders in England. *Social Psychiatry and Psychiatric Epidemiology*, *48*(5), 685–692. https://doi.org/10.1007/s00127-012-0565-y

Costello, E. J., Egger, H. L., & Angold, A. (2005). The developmental epidemiology of anxiety disorders: Phenomenology, prevalence, and comorbidity. *Child and Adolescent*

Psychiatric Clinics of North America, *14*(4), 631–648, vii. https://doi.org/10.1016/j.chc.2005.06.003

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, *43*(6), 1241–1299. https://doi.org/10.2307/1229039

Department of Census and Statistics Sri Lanka. (2019). *Sri Lanka Labour Force Survey—Annual Report 2019* [Annual Report]. Ministry of Finance. http://www.statistics.gov.lk/LabourForce/StaticalInformation/AnnualReports/2019

Department of Census and Statistics Sri Lanka. (2022). *Economic Statistics of Sri Lanka.* http://www.statistics.gov.lk/Publication/Economic-Statistics-2022

Duncan, G. J., & Magnuson, K. A. (2003). Off with Hollingshead: Socioeconomic resources, parenting, and child development. In *Socioeconomic status, parenting, and child development* (pp. 83–106). Lawrence Erlbaum Associates Publishers.

Eilertsen, E. M., Jami, E. S., McAdams, T. A., Hannigan, L. J., Havdahl, A. S., Magnus, P., Evans, D. M., & Ystrom, E. (2021). Direct and Indirect Effects of Maternal, Paternal, and Offspring Genotypes: Trio-GCTA. *Behavior Genetics*. https://doi.org/10.1007/s10519-020-10036-6

Festin, K., Thomas, K., Ekberg, J., & Kristenson, M. (2017). Choice of measure matters: A study of the relationship between socioeconomic status and psychosocial resources in a middle-aged normal population. *PLOS ONE*, *12*(8), e0178929. https://doi.org/10.1371/journal.pone.0178929

Greenwood, J., Guner, N., Kocharkov, G., & Santos, C. (2014). Marry Your Like: Assortative Mating and Income Inequality. *American Economic Review*, *104*(5), 348–353. https://doi.org/10.1257/aer.104.5.348

Hannigan, L. J., Walaker, N., Waszczuk, M. A., McAdams, T. A., & Eley, T. C. (2017). Aetiological Influences on Stability and Change in Emotional and Behavioural Problems across Development: A Systematic Review. *Psychopathology Review*, *a4*(1), 52–108. https://doi.org/10.5127/pr.038315

Hardy, B., & Ziliak, J. P. (2014). Decomposing Trends in Income Volatility: The "Wild Ride" at the Top and Bottom. *Economic Inquiry*, *52*(1), 459–476. https://doi.org/10.1111/ecin.12044

Harold, G. T., & Conger, R. D. (1997). Marital conflict and adolescent distress: The role of adolescent awareness. *Child Development*, *68*(2), 333–350. https://doi.org/10.1111/j.1467-8624.1997.tb01943.x

Hendriks, A. M., Finkenauer, C., Nivard, M. G., Van Beijsterveldt, C. E. M., Plomin, R. J., Boomsma, D. I., & Bartels, M. (2019). Comparing the genetic architecture of childhood behavioral problems across socioeconomic strata in the Netherlands and the United Kingdom. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-019-01357-x

Hill, H. D., Morris, P., Gennetian, L. A., Wolf, S., & Tubbs, C. (2013). The Consequences of Income Instability for Children's Well-Being. *Child Development Perspectives*, *7*(2), 85–90. https://doi.org/10.1111/cdep.12018

Howe, L. D., Galobardes, B., Matijasevich, A., Gordon, D., Johnston, D., Onwujekwe, O., Patel, R., Webb, E. A., Lawlor, D. A., & Hargreaves, J. R. (2012). Measuring socio-economic position for epidemiological studies in low- and middle-income countries: A methods of measurement in epidemiology paper. *International Journal of Epidemiology*, *41*(3), 871–886. https://doi.org/10.1093/ije/dys037

Howe, L. D., Tilling, K., Galobardes, B., & Lawlor, D. A. (2013). Loss to follow-up in cohort studies: Bias in estimates of socioeconomic inequalities. *Epidemiology (Cambridge, Mass.)*, *24*(1), 1–9. https://doi.org/10.1097/EDE.0b013e31827623b1

Hur, Y.-M. (2008). Sex Differences in Genetic and Environmental Contributions to Depression Symptoms in South Korean Adolescent and Young Adult Twins. *Twin Research and Human Genetics*, *11*(3), 306–313. https://doi.org/10.1375/twin.11.3.306

Jacobs, E., & Hacker, J. (2008). *The Rising Instability of American Family Incomes, 1969-2004: Evidence from the Panel Study of Income Dynamics*. https://policycommons.net/artifacts/1414804/the-rising-instability-of-american-family-incomes-1969-2004/2029067/

Jacobsen, T. N., Nohr, E. A., & Frydenberg, M. (2010). Selection by socioeconomic factors into the Danish National Birth Cohort. *European Journal of Epidemiology*, *25*(5), 349–355. https://doi.org/10.1007/s10654-010-9448-2

Jayaweera, K., Aschan, L., Pannala, G., Adikari, A., Glozier, N., Ismail, K., Pariante, C. M., Rijsdijk, F., Siribaddana, S., Zavos, H. M. S., Zunszain, P. A., Sumathipala, A., & Hotopf,

M. (2018). The Colombo Twin and Singleton Follow-up Study: A population based twin study of psychiatric disorders and metabolic syndrome in Sri Lanka. *BMC Public Health*, *18*(1), 145. https://doi.org/10.1186/s12889-017-4992-2

Jin, Y., Zhu, D., & He, P. (2020). Social causation or social selection? The longitudinal interrelationship between poverty and depressive symptoms in China. *Social Science & Medicine*, *249*, 112848. https://doi.org/10.1016/j.socscimed.2020.112848

Keers, R., Coleman, J. R. I., Lester, K. J., Roberts, S., Breen, G., Thastum, M., Bögels, S., Schneider, S., Heiervang, E., Meiser-Stedman, R., Nauta, M., Creswell, C., Thirlwall, K., Rapee, R. M., Hudson, J. L., Lewis, C., Plomin, R., & Eley, T. C. (2016). A Genome-Wide Test of the Differential Susceptibility Hypothesis Reveals a Genetic Predictor of Differential Response to Psychological Treatments for Child Anxiety Disorders. *Psychotherapy and Psychosomatics*, *85*(3), 146–158. https://doi.org/10.1159/000444023

Kola, L., Kohrt, B. A., Hanlon, C., Naslund, J. A., Sikander, S., Balaji, M., Benjet, C., Cheung, E. Y. L., Eaton, J., Gonsalves, P., Hailemariam, M., Luitel, N. P., Machado, D. B., Misganaw, E., Omigbodun, O., Roberts, T., Salisbury, T. T., Shidhaye, R., Sunkel, C., … Patel, V. (2021). COVID-19 mental health impact and responses in low-income and middle-income countries: Reimagining global mental health. *The Lancet Psychiatry*, *8*(6), 535–550. https://doi.org/10.1016/S2215-0366(21)00025-0

Lorant, V., Demarest, S., Miermans, P.-J., & Van Oyen, H. (2007). Survey error in measuring socio-economic risk factors of health status: A comparison of a survey and a census. *International Journal of Epidemiology*, *36*(6), 1292–1299. https://doi.org/10.1093/ije/dym191

Lund, C., & Cois, A. (2018). Simultaneous social causation and social drift: Longitudinal analysis of depression and poverty in South Africa. *Journal of Affective Disorders*, *229*, 396–402. https://doi.org/10.1016/j.jad.2017.12.050

Macintyre, S., Maciver, S., & Sooman, A. (1993). Area, Class and Health: Should we be Focusing on Places or People? *Journal of Social Policy*, *22*(2), 213–234. https://doi.org/10.1017/S0047279400019310

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

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature*, *456*(7218), Article 7218. https://doi.org/10.1038/456018a

Mann, E. M., Heesch, K. C., Rachele, J. N., Burton, N. W., & Turrell, G. (2022). Individual socioeconomic position, neighbourhood disadvantage and mental well-being: A cross-sectional multilevel analysis of mid-age adults. *BMC Public Health*, *22*(1), 494. https://doi.org/10.1186/s12889-022-12905-7

Mascayano, F., Armijo, J. E., & Yang, L. H. (2015). Addressing Stigma Relating to Mental Illness in Low- and Middle-Income Countries. *Frontiers in Psychiatry*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyt.2015.00038

Maselko, J. (2017). Social Epidemiology and Global Mental Health: Expanding the Evidence from High-Income to Low- and Middle-Income Countries. *Current Epidemiology Reports*, *4*(2), 166–173. https://doi.org/10.1007/s40471-017-0107-y

Matthias, A. T., & Jayasinghe, S. (2022). Worsening economic crisis in Sri Lanka: Impacts on health. *The Lancet Global Health*, *10*(7), e959. https://doi.org/10.1016/S2214-109X(22)00234-0

Maura, J., & Weisman de Mamani, A. (2017). Mental Health Disparities, Treatment Engagement, and Attrition Among Racial/Ethnic Minorities with Severe Mental Illness: A Review. *Journal of Clinical Psychology in Medical Settings*, *24*(3–4), 187–210. https://doi.org/10.1007/s10880-017-9510-2

Metten, M.-A., Costet, N., Multigner, L., Viel, J.-F., & Chauvet, G. (2022). Inverse probability weighting to handle attrition in cohort studies: Some guidance and a call for caution. *BMC Medical Research Methodology*, *22*(1), 45. https://doi.org/10.1186/s12874-022-01533-9

Middeldorp, C. M., Lamb, D. J., Vink, J. M., Bartels, M., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2014). Child care, socio-economic status and problem behavior: A study of gene-environment interaction in young Dutch twins. *Behavior Genetics*, *44*(4), 314–325. https://doi.org/10.1007/s10519-014-9660-z

Miller, P., Betancur, L., Whitfield, K., & Votruba-Drzal, E. (2021). Examining income dynamics and externalizing and internalizing trajectories through a developmental

psychopathology lens: A nationally representative study. *Development and Psychopathology*, *33*(1), 1–17. https://doi.org/10.1017/S0954579419001494

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

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, *110*, 406–425. https://doi.org/10.1037/0033-2909.110.3.406

Morrissey, T. W., Cha, Y., Wolf, S., & Khan, M. (2020). Household economic instability: Constructs, measurement, and implications. *Children and Youth Services Review*, *118*, 105502. https://doi.org/10.1016/j.childyouth.2020.105502

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: When selection bias can substantially influence observed associations. *International Journal of Epidemiology*, *47*(1), 226–235. https://doi.org/10.1093/ije/dyx206

Murray, A. L., Molenaar, D., Johnson, W., & Krueger, R. F. (2016). Dependence of Gene-by-Environment Interactions (GxE) on Scaling: Comparing the Use of Sum Scores, Transformed Sum Scores and IRT Scores for the Phenotype in Tests of GxE. *Behavior Genetics*, *46*(4), 552–572. https://doi.org/10.1007/s10519-016-9783-5

National Institute for Health Research. (2023). *Patient and Public Involvement (PPI)*. Research Design Service South Central. https://www.rds-sc.nihr.ac.uk/ppi-information-resources/

Nohr, E. A., & Liew, Z. (2018). How to investigate and adjust for selection bias in cohort studies. *Acta Obstetricia et Gynecologica Scandinavica*, *97*(4), 407–416. https://doi.org/10.1111/aogs.13319

Paré, G., Cook, N. R., Ridker, P. M., & Chasman, D. I. (2010). On the use of variance per genotype as a tool to identify quantitative trait interaction effects: A report from the Women's Genome Health Study. *PLoS Genetics*, *6*(6), e1000981. https://doi.org/10.1371/journal.pgen.1000981

Plomin, R. (2022). The next 10 years of behavioural genomic research. *JCPP Advances*, *2*(4), e12112. https://doi.org/10.1002/jcv2.12112

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), Article 7. https://doi.org/10.1038/ng.3285

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

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119–133. https://doi.org/10.1093/bib/3.2.119

Rimfeld, K., Krapohl, E., Trzaskowski, M., Coleman, J. R. I., Selzam, S., Dale, P. S., Esko, T., Metspalu, A., & Plomin, R. (2018). Genetic influence on social outcomes during and after the Soviet era in Estonia. *Nature Human Behaviour*, *2*(4), 269–275. https://doi.org/10.1038/s41562-018-0332-5

Samarasekara, N., Davies, M. L. M., & Siribaddana, S. (2012). The stigma of mental illness in Sri Lanka: The perspectives of community mental health workers. *Stigma Research and Action*, *2*(2). https://doi.org/10.5463/sra.v1i3.48

Santomauro, D. F., Herrera, A. M. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., Abbafati, C., Adolph, C., Amlag, J. O., Aravkin, A. Y., Bang-Jensen, B. L., Bertolacci, G. J., Bloom, S. S., Castellano, R., Castro, E., Chakrabarti, S., Chattopadhyay, J., Cogen, R. M., Collins, J. K., … Ferrari, A. J. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*, *398*(10312), 1700–1712. https://doi.org/10.1016/S0140-6736(21)02143-7

Scaramella, L. V., Conger, R. D., & Simons, R. L. (1999). Parental protective influences and gender-specific increases in adolescent internalizing and externalizing problems. *Journal of Research on Adolescence*, *9*(2), 111–141. https://doi.org/10.1207/s15327795jra0902_1

Schmalbach, B., Zenger, M., Tibubos, A. N., Kliem, S., Petrowski, K., & Brähler, E. (2021). Psychometric Properties of Two Brief Versions of the Hopkins Symptom Checklist: HSCL-5 and HSCL-10. *Assessment*, *28*(2), 617–631. https://doi.org/10.1177/1073191119860910

Siribaddana, S. H., Ball, H. A., Hewage, S. N., Glozier, N., Kovas, Y., Dayaratne, D., Sumathipala, A., McGuffin, P., & Hotopf, M. (2008). Colombo Twin and Singleton Study (CoTASS): A description of a population based twin study of mental disorders in Sri Lanka. *BMC Psychiatry*, *8*(1), 49. https://doi.org/10.1186/1471-244X-8-49

Stafford, M., & Marmot, M. (2003). Neighbourhood deprivation and health: Does it affect us all equally? *International Journal of Epidemiology*, *32*(3), 357–366. https://doi.org/10.1093/ije/dyg084

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

Torvik, F. A., Eilertsen, E. M., McAdams, T. A., Gustavson, K., Zachrisson, H. D., Brandlistuen, R., Gjerde, L. C., Havdahl, A., Stoltenberg, C., Ask, H., & Ystrom, E. (2020). Mechanisms linking parental educational attainment with child ADHD, depression, and academic problems: A study of extended families in The Norwegian Mother, Father and Child Cohort Study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. https://doi.org/10.1111/jcpp.13197

Torvik, F. A., Rosenström, T. H., Gustavson, K., Ystrom, E., Kendler, K. S., Bramness, J. G., Czajkowski, N., & Reichborn-Kjennerud, T. (2019). Explaining the association between anxiety disorders and alcohol use disorder: A twin study. *Depression and Anxiety*, *36*(6), 522–532. https://doi.org/10.1002/da.22886

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socioeconomic status: Gene-environment interaction. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(7), 734–743. https://doi.org/10.1111/j.1469-7610.2005.01552.x

United Nations. (2020). *World Social Report 2020: Inequality in a Rapidly Changing World*. UN. https://doi.org/10.18356/7f5d0efc-en

Van Hulle, C. A., & Rathouz, P. J. (2015). Operating Characteristics of Statistical Methods for Detecting Gene-by-Measured Environment Interaction in the Presence of Gene-Environment Correlation under Violations of Distributional Assumptions. *Twin*

Research and Human Genetics: The Official Journal of the International Society for Twin Studies, *18*(1), 19–27. https://doi.org/10.1017/thg.2014.81

Visscher, P. M., & Posthuma, D. (2010). Statistical power to detect genetic Loci affecting environmental sensitivity. *Behavior Genetics*, *40*(5), 728–733. https://doi.org/10.1007/s10519-010-9362-0

Visser, K., Bolt, G., Finkenauer, C., Jonker, M., Weinberg, D., & Stevens, G. W. J. M. (2021). Neighbourhood deprivation effects on young people's mental health and well-being: A systematic review of the literature. *Social Science & Medicine*, *270*, 113542. https://doi.org/10.1016/j.socscimed.2020.113542

Whitfield, K., Betancur, L., Miller, P., & Votruba-Drzal, E. (2021). Longitudinal Links between Income Dynamics and Young Adult Socioeconomic and Behavioral Health Outcomes. *Youth & Society*, *53*(7), 1181–1210. https://doi.org/10.1177/0044118X21996382

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A Tool for Genome-wide Complex Trait Analysis. *American Journal of Human Genetics*, *88*(1), 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Zavos, H. M. S., Dalton, B., Jayaweera, K., Harber-Aschan, L., Pannala, G., Adikari, A., Hatch, S. L., Siribaddana, S., Sumathipala, A., Hotopf, M., & Rijsdijk, F. V. (2020). The relationship between independent and dependent life events and depression symptoms in Sri Lanka: A twin and singleton study. *Social Psychiatry and Psychiatric Epidemiology*, *55*(2), 237–249. https://doi.org/10.1007/s00127-019-01765-z

Appendix A

Supplementary materials for Chapter 2

Table S1. Phenotypic correlations between all study variables.

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10
1.Maternal age	30.30	4.49										
2. Paternal age	32.28	5.26	.67** [.66, .67]									
3. Parity	0.78	0.88	.40** [.40, .41]	.32** [.31, .32]								
4. Child year of birth	2005.07	2.20	.07** [.06, .08]	.05** [.04, .06]	-.07** [-.08, -.07]							
5. Child sex	1.49	0.50	.00 [-.01, .01]	-.00 [-.01, .01]	-.00 [-.01, .01]	-.00 [-.01, .01]						
6. Child emotional problems	1.25	0.20	-.08** [-.09, -.08]	-.04** [-.05, -.03]	-.08** [-.08, -.07]	-.02** [-.03, -.02]	.03** [.02, .04]					
7. Child behavioural problems	1.47	0.26	-.09** [-.09, -.08]	-.05** [-.05, -.04]	-.06** [-.07, -.05]	-.08** [-.09, -.08]	-.08** [-.08, -.07]	.39** [.38, .39]				
8. Maternal registry income	0.57	0.24	.06** [.06, .07]	.02** [.01, .02]	-.10** [-.10, -.09]	.09** [.08, .10]	.01 [-.00, .01]	-.08** [-.09, -.08]	-.05** [-.05, -.04]			
9. Maternal registry education	5.37	1.38	.22** [.21, .22]	.11** [.10, .12]	-.09** [-.10, -.09]	.15** [.14, .16]	.01 [-.00, .01]	-.14** [-.14, -.13]	-.14** [-.14, -.13]	.38** [.37, .38]		
10. Paternal registry income	0.61	0.23	.05** [.05, .06]	-.04** [-.04, -.03]	-.01** [-.02, -.00]	.04** [.03, .05]	.00 [-.01, .01]	-.06** [-.06, -.05]	-.05** [-.06, -.05]	.07** [.06, .08]	.16** [.15, .17]	
11. Paternal registry education	5.06	1.54	.20** [.19, .21]	.08** [.07, .09]	-.03** [-.03, -.02]	.10** [.10, .11]	.00 [-.01, .01]	-.10** [-.11, -.09]	-.11** [-.11, -.10]	.19** [.18, .20]	.45** [.44, .46]	.33** [.33, .34]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates $p < .05$. ** indicates $p < .01$.

Table S2. Fit statistics from the biometric moderation MCoTS models of child emotional and behavioural outcomes moderated by maternal SES indices.

	Δ -2LL	Δ *df*	*p*	AIC
Maternal income and child emotional problems				
Full moderation	-	-	-	188201.00
No moderation	85.78	6	2.28e-16	188274.80
Maternal education and child emotional problems				
Full moderation				175788.00
No moderation	155.74	6	4.73e-31	175931.70
Maternal income and child behavioural problems				
Full moderation	-	-	-	188551.30
No moderation	12.57	6	0.05	188551.90
Maternal education and child behavioural problems				
Full moderation	-	-	-	176167.40
No moderation	45.32	6	4.05e-08	176200.70

-2LL = -2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion.

Table S3. Moderated path estimates, standard errors and 95% confidence intervals of child emotional and behavioural outcomes moderated by maternal SES indices.

Child emotional problems								
	Maternal income				Maternal education			
Parameter	Estimate	Std. error	LL	UL	Estimate	Std. error	LL	UL
β_{xu}	-0.06	0.06	-0.14	0.14	0.11	0.04	-0.17	0.17
β_{yu}	-0.03	0.03	-0.08	0.03	**-0.06**	**0.02**	**-0.11**	**-0.01**
β_{xc}	0.02	0.05	-0.07	0.11	0.05	0.04	-0.04	0.12
β_{zu}	0.01	0.03	-0.05	0.06	0.01	0.02	-0.03	0.05
β_{yc}	-0.09	0.03	-0.14	0.14	-0.04	0.07	-0.16	0.16
β_{zc}	-0.01	0.02	-0.04	0.03	-0.03	0.02	-0.06	0.01

Child behavioural problems								
	Maternal income				Maternal education			
Parameter	Estimate	Std. error	LL	UL	Estimate	Std. error	LL	UL
β_{xu}	0.00	0.05	-0.09	0.09	0.00	0.23	-0.17	0.17
β_{yu}	0.00	0.01	-0.03	0.04	-0.03	0.07	-0.08	0.05
β_{xc}	**0.06**	**0.02**	**0.01**	**0.09**	0.08	0.14	-0.03	0.14
β_{zu}	**0.06**	**0.02**	**0.02**	**0.09**	**0.06**	**0.03**	**0.02**	**0.08**
β_{yc}	0.02	0.03	-0.03	0.08	0.09	0.06	-0.15	0.15
β_{zc}	-0.01	0.01	-0.03	0.01	-0.04	0.04	-0.07	0.01

Note: Moderated components estimated under the full moderation model. The parameter estimates for the mean and unmoderated parameters are not shown. β_{xu}, β_{yu} and β_{zu} = moderated components of A2, C2 and E2 (i.e. variance components unique to child emotional or behavioural traits). β_{xc}, β_{yc} and β_{zc} = moderated components of A1', C1' and p (i.e. variance components common to parent SES and child emotional or behavioural traits). LL = lower bound of the 95% confidence interval. UL = upper bound of the 95% confidence interval.

Table S4. Fit statistics from the biometric moderation MCoTS models of child emotional and behavioural outcomes moderated by paternal SES indices.

	Δ -2LL	Δ *df*	*p*	AIC
Paternal income and child emotional problems				
Full moderation	-	-	-	185850.90
No moderation	59.98	6	4.55e-11	185898.90
Paternal education and child emotional problems				
Full moderation	-	-	-	170848.20
No moderation	83.53	6	6.64e-16	170919.80
Paternal income and child behavioural problems				
Full moderation	-	-	-	186114.00
No moderation	4.38	6	0.63	186106.40
Paternal education and child behavioural problems				
Full moderation	-	-	-	171142.60
No moderation	10.01	6	0.12	171140.60

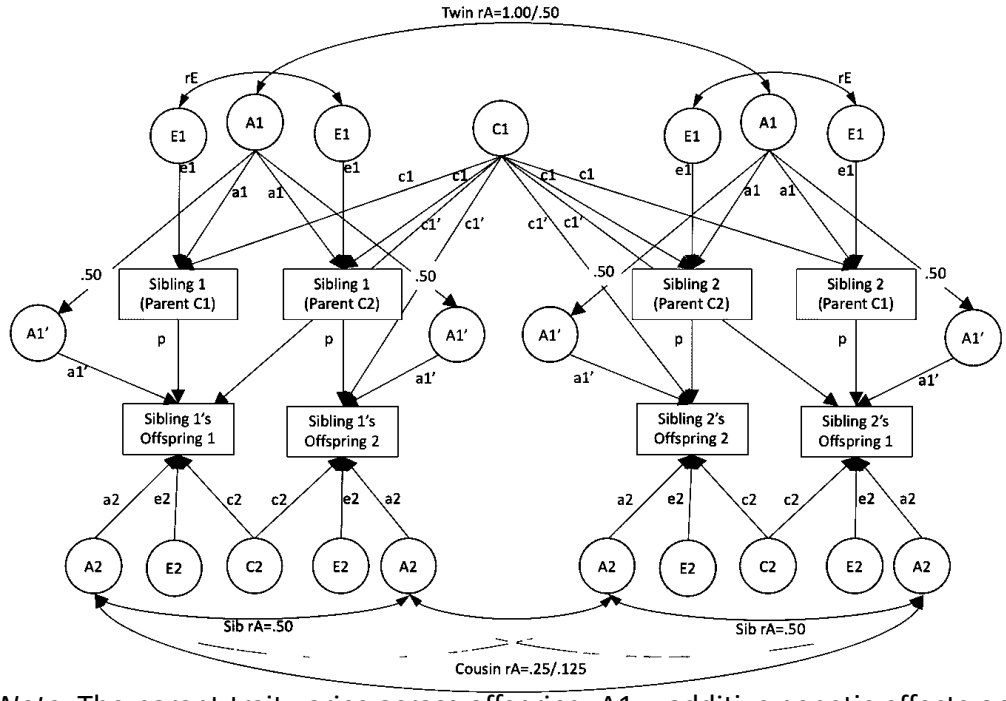
-2LL = -2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion.

Table S5. Moderated path estimates, standard errors and 95% confidence intervals of child emotional and behavioural outcomes moderated by paternal SES indices.

Child emotional problems								
	Paternal income				Paternal education			
Parameter	Estimate	Std. error	LL	UL	Estimate	Std. error	LL	UL
β_{xu}	0.04	0.05	-0.13	0.13	0.00	0.08	-0.10	0.10
β_{yu}	-0.06	0.04	-0.12	0.07	**-0.04**	**0.01**	**-0.06**	**-0.02**
β_{xc}	0.03	0.09	-0.12	0.14	0.01	0.03	-0.06	0.07
β_{zu}	-0.04	0.03	-0.08	0.04	-0.02	0.03	-0.06	0.03
β_{yc}	-0.08	0.06	-0.15	0.15	0.01	0.04	-0.07	0.08
β_{zc}	0.00	0.03	-0.05	0.05	0.00	0.02	-0.04	0.03
Child behavioural problems								
	Paternal income				Paternal education			
Parameter	Estimate	Std. error	LL	UL	Estimate	Std. error	LL	UL
β_{xu}	-0.02	0.03	-0.08	0.06	0.00	0.06	-0.10	0.10
β_{yu}	0.03	0.02	-0.02	0.08	-0.01	0.03	-0.06	0.03
β_{xc}	-0.03	0.04	-0.10	0.05	0.04	0.04	-0.06	0.11
β_{zu}	-0.02	0.02	-0.06	0.02	-0.04	0.02	NA	0.02
β_{yc}	-0.04	0.04	-0.10	0.10	0.00	0.03	-0.06	0.07
β_{zc}	0.01	0.01	-0.02	0.04	-0.02	0.02	NA	0.03

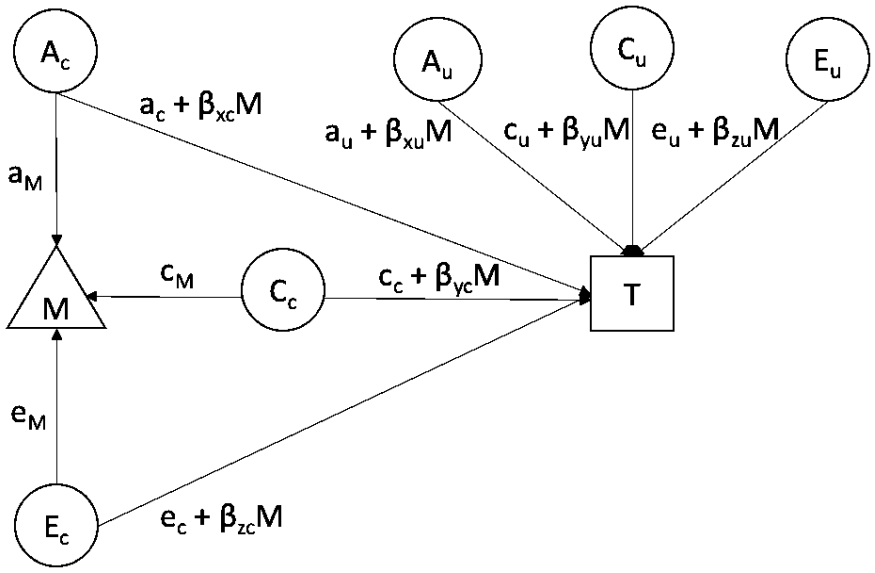
Note: Moderated components estimated under the full moderation model. The parameter estimates for the mean and unmoderated parameters are not shown. β_{xu}, β_{yu} and β_{zu} = moderated components of A2, C2 and E2 (i.e. variance components unique to child emotional or behavioural traits). β_{xc}, β_{yc} and β_{zc} = moderated components of A1', C1' and p (i.e. variance components common to parent SES and child emotional or behavioural traits). LL = lower bound of the 95% confidence interval. UL = upper bound of the 95% confidence interval.

Figure S1. Path diagram of the full Multiple-Children-of-Twins-and-Siblings (MCoTS) structural equation model.



Note. The parent trait varies across offspring. A1 = additive genetic effects on parental trait; C1 = shared environmental effects on parental trait; E1 = nonshared environmental effects on parental trait; A1' = genetic effects shared between parental trait and offspring trait; C1' = extended family effects (i.e. shared environment of the parents influences offspring trait); A2 = genetic effects specific to offspring trait; C2 = shared environmental effects on offspring trait; E2 = nonshared environmental effects on offspring trait; p = residual phenotypic association after accounting for genetic and environmental overlap; rE = within-parent correlation between E1 for parent trait of child 1 and 2. The path between A1 and A1' is fixed to 0.5 because parents share half their DNA with their children.

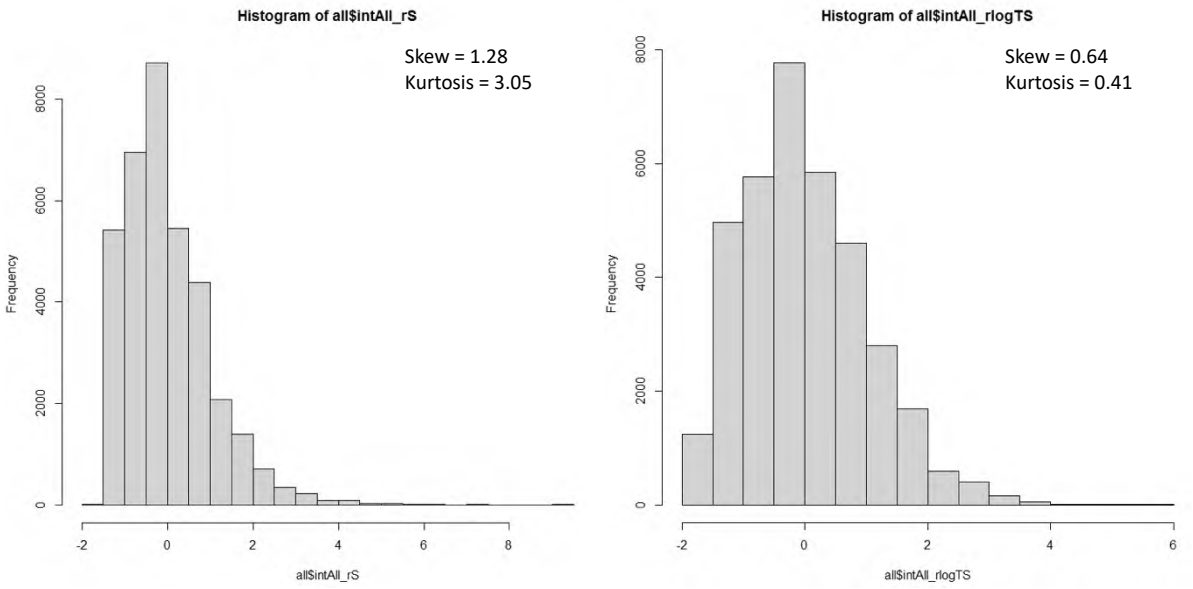
Figure S2. Purcell (2002) bivariate moderation model.



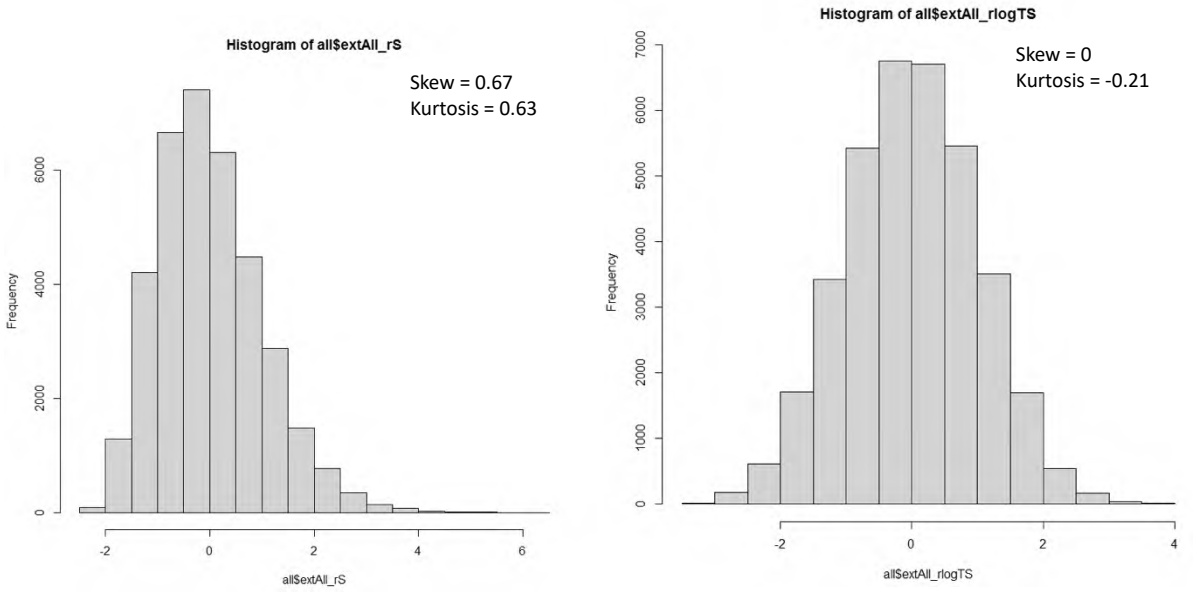
Note: Bivariate moderation model shown for only one member of a twin pair as proposed by Purcell (2002). Ac, Cc and Ec are the variance components common to the moderator and the trait. Au, Cu, and Eu are the variance components unique to the trait. β coefficients index the direction and magnitude of moderation. The total variance of the trait can be calculated as follows: $Var(T|M) = (a_c + \beta_{xc}M)^2 + (a_u + \beta_{xu}M)^2 + (c_c + \beta_{yc}M)^2 + (c_u + \beta_{yu}M)^2 + (e_c + \beta_{zc}M)^2 + (e_u + \beta_{zu}M)^2$.

Figure S3. Distributions of child emotional and behavioural outcome variables before and after logarithmic transformation.

(a)



(b)



Note: **(a)** Histograms of emotional scores scale before (left) and after log transformation (right). **(b)** Histograms of behavioural scores scale before (left) and after log transformation (right).

Appendix B

Supplementary materials for Chapter 3

Table S1. Items included in standard of living composite.

Items	**Response options**	
Housing composite		
Tenure of dwelling	Owned by employer/government; Rent free (relative/friend/employer); Rent/Lease; Other	0
	Owned by the occupant (/member in the immediate family)	1
Type of structure	Line Room/Row house; Slum/Shanty; Other	0
	Single house; Condo; Attached house/Annex/Flat	1
Rooms used for sleeping	One room	0
	More than one room	1
Main material of the floor	Wood planks/Bamboo; /Mud/Clay/Sand; Dung; Other	0
	Cement; Terrazzo/Tile/Granite; Carpet; Vinyl; Polished wood	1
Main material of the roof	Metal Sheet; Cadjan/Palmyrah/Straw; Other	0
	Tile; Asbestos Sheets; Concrete	1
Main material of the wall	Brick; Cement block; Cabook; Granite	0
	Mud/Mud block; Cadjan; Palmyrah; Plank; Metal sheet; Wattle and Daub; Other	1
Principal type of lighting	Kerosene/Solar/Battery/Prashakthi; Other; None	0
	Electricity	1
Principal type of cooking fuel	Firewood/Kerosene; Saw dust; Paddy husk; Other	0
	Gas	1
Source of drinking water	Public tap/Street tap; Protected well outside premises/ Rainwater Collection; River; Tank; Streams; Other	0
	Tap in the yard or premises; Tube well; Protected well within the premises	1
	Tap within premises	2
Ownership of household appliances		
Does your house have a	No/Yes	0/1
Radio		
Tv		
Refrigerator		
Phone (mobile/fixed)		
Computer/Laptop		
Internet access		
Access to any transport	No/Yes	0/1

Table S2. Unadjusted associations between sociodemographic characteristics and socioeconomic status indicators.

		Standard of living	**Educational attainment**	**Financial stability**
	N	Unadjusted β	Unadjusted β	Unadjusted β
Sex				
Male (ref)	1681 (42.4%)			
Female	2288 (57.6%)	-.16 [-.23, -.09]**	-.02 [-.09, .05]	-.17 [-.24, -.10]**
Age				
19-29 (ref)	853 (21.5%)			
30-39	1012 (25.5%)	-.24 [-.36,-.13]**	-.45 [-.56, .34]**	-.15 [-.25, -.06]**
40-49	825 (20.8%)	-.21 [-.32, -.13]**	-.62 [-.73, -.51]**	-.31 [-.42, -.20]**
50-59	665 (16.8%)	-.17 [-.28, -.05]**	-.74 [-.85, -.62]**	-.28 [-.39, -.17]**
60-69	376 (9.5%)	-.25 [-.39, -.11]**	-.83 [-.96, -.70]**	-.37 [-.50, -.23]**
>70	203 (5.1%)	-.25 [-.44, -.06]*	-.92 [-1.09, -.75]**	-.14 [-.32, .04]
Ethnicity				
Sinhala (ref)	3647 (91.9%)			
Tamil	120 (3.0%)	-.45 [-.75, -.16]**	-.37 [-.63, -.12]**	-.29 [-.53, -.05]*
Muslim	150 (3.8%)	.12 [-.07, .32]	-.39 [-.55, -.24]**	-.06 [-.23, .12]
Other Minority	16 (0.4%)	.17 [-.06, .41]	-.09 [-.55, .38]	-.34 [-.84, .15]
Marital Status				
Married (ref)	2838 (71.5%)			
Previously Married	329 (8.3%)	-.32 [-.42, -.23]**	-.51 [-.63, -.40]**	-.30 [-.44, -.16]**
Never Married	763 (19.2%)	.18 [.09, .28]**	.49 [.40, .59]**	.17 [.08, .25]**
Urbanicity				
Urban (ref)	2390 (60.2%)			
Rural	532 (13.4%)	-.32 [-.42, -.19]**	-.13 [-.23, -.02]*	.18 [.11, .26]**

Mixed	826 (20.8%)	-.17 [-.26, -.08]**	-.15 [-.24, -.07]**	.16 [.07, .24]**
Outside Colombo	221 (5.6%)	-.23 [-.36, -.09]**	.08 [-.06, .21]	-.10 [-.24, .05]
Standard of Living				
Mean (SD)	14.1 (2.64)	-	.18 [.17, .19]**	.15 [.14, .17]**
Educational attainment				
No education (ref)	47 (1.2%)		-	
Grade 1-5	274 (6.9%)	.13 [-.31, .57]	-	-.15 [-.59, .29]
Grade 6 0/Ls	1757 (44.3%)	.57 [.15, 1.00]**	-	.30 [-.12, .72]
Passed O/Ls	632 (15.9%)	1.16 [.73, 1.59]**	-	.72 [.30, 1.14]**
Up to/ passed A/Ls	929 (23.4%)	1.44 [1.01, 1.87**]	-	.87 [.45, 1.28]**
University /higher	276 (7.0%)	1.79 [1.36, 2.23]**	-	1.00 [.58, 1.43]**
Financial Stability				-
Very difficult to make ends meet (ref)	121 (3.0%)			-
Difficult to make ends meet	284 (7.2%)	.64 [.37, .91]**	.34 [.16, .51]**	-
Just about getting by	547 (13.8%)	.80 [.54, 1.05]**	.41 [.25, .57]**	-
Doing alright	2616 (65.9%)	1.36 [1.12, 1.61]**	.92 [.77, 1.07]**	-
Living comfortably	365 (9.2%)	2.10 [1.84, 2.35]**	1.15 [1.33, 1.68]**	-

Note. Linear regressions were conducted using standardised outcome variables and clustered standard errors to account for non-independence of twins in the sample.
*p < 0·05; **p < 0·01

Table S3. Unadjusted and adjusted (for sex and age) associations between depression symptoms and socioeconomic status indicators.

	Mean (SD)	Standard of living β (95% CI)	Education β (95% CI)	Financial stability β (95% CI)
Depression symptoms	4.86 (6.19)			
Unadjusted		-.14 [-.17, -.10]**	-.13 [-.17, -.10]**	-.14 [-.18, -.10]**
Adjusted model[a]		-.13 [-.16, -.09]**	-.12 [-.15, -.08]**	-.12 [-.16, -.08]**

Note. Standardised depression and SES indicators were analysed as continuous variables using linear regression. [a] Adjusted for sex and age.
**p < 0.01

Table S4. Fit statistics for univariate ACE models for standard of living, educational attainment, financial stability, and depression symptoms.

Measure	Model	Model Fit					
		-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Standard of living	Constrained correlation	10406.71	3846	10424.71	-	-	-
	Heterogeneity	**10407.09**	**3847**	**10423.09**	**0.39**	**1**	**0.535**
	Scalar	10414.30	3849	10426.30	7.21	2	0.027
	Homogeneity	10418.93	3850	10428.93	11.84	3	0.008
Educational attainment	Constrained correlation	10322.37	3831	10340.37	-	-	-
	Heterogeneity	10323.91	3832	10339.91	1.54	1	0.215
	Scalar	10328.96	3834	10340.96	5.06	2	0.080
	Homogeneity	**10329.88**	**3835**	**10339.88**	**5.97**	**3**	**0.113**
Financial stability	Constrained correlation	10648.05	3848	10666.05	-	-	-
	Heterogeneity	10664.85	3849	10680.85	16.80	1	4.15e-05
	Scalar	10670.06	3851	10682.06	5.21	2	0.074
	Homogeneity	10695.57	3852	10705.57	30.72	3	9.75e-07
Depression	Constrained correlation	10638.44	3830	10656.44	-	-	-
	Heterogeneity	10638.60	3831	10654.60	0.16	1	0.690
	Scalar	**10641.48**	**3833**	**10653.48**	**2.88**	**2**	**0.237**
	Homogeneity	10790.42	3834	10800.42	151.82	3	1.07e-32

Note. Constrained correlation model: the means and standard deviation equated across birth-order. Heterogeneity model: sex differences in the magnitude of genetic and environmental influences on the trait. Scalar model: genetic and environmental influences are equated across males and females but variance differences across sexes are modelled. Homogeneity model: genetic and environmental influences are equated across males and females. -2LL = negative 2* log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion (lower values reflect a better fit); Δ -2LL = likelihood ratio chi-square (X^2) test comparing the difference in -2LL of the models; Δ *df* = difference in degrees of freedom of the models; p = *p*-value of the X^2 test.

Table S5. Fit statistics from the no-sex-differences bivariate moderation models of depression symptoms moderated by standard of living and educational attainment.

	-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Depression moderated by standard of living						
Full moderation	18861.11	6880	18903.11	-	-	-
No moderation	18876.78	6886	18906.78	15.67	6	0.016
Drop βau, βcu and βeu	18873.56	6883	18909.56	12.45	3	0.006
Drop βac, βcc and βec	18862.24	6883	18898.24	1.13	3	0.770
Depression moderated by educational attainment						
Full moderation	18637.07	6830	18679.07	-	-	-
No moderation	18657.97	6836	18687.97	20.90	6	0.002
Drop βau, βcu and βeu	18655.13	6833	18691.13	18.06	3	4.28e-04
Drop βac, βcc and βec	18640.57	6833	18676.57	3.50	3	0.321

Note. βau, βcu, and βeu are the moderated variance components unique to depression symptoms. βac, βcc, and βec are the moderated variance components common to the moderator (i.e., standard of living or educational attainment) and depression symptoms. -2LL = negative 2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion (lower values reflect a better fit); Δ -2LL = likelihood ratio chi-square (X^2) test comparing the difference in -2LL of the models; Δ *df* = difference in degrees of freedom of the models; p = *p*-value of the X^2 test.

Table S6. Fit statistics from the no-sex-differences bivariate phenotypic moderation models of depression symptoms moderated by standard of living and educational attainment.

	-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Depression moderated by standard of living						
Full moderation	18808.36	6886	18838.36	-	-	-
No moderation	18821.59	6888	18847.59	13.24	2	0.001
Drop βu	18819.31	6887	18847.31	10.95	1	0.001
Drop βc	18810.02	6887	18838.02	1.66	1	0.197
Depression moderated by educational attainment						
Full moderation	18587.34	6836	18617.34	-	-	-
No moderation	18603.51	6838	18629.51	16.17	2	3.08e-04
Drop βu	18603.13	6837	18631.13	15.78	1	7.10e-05
Drop βc	18587.69	6837	18615.69	0.35	1	0.55

Note. βu is the moderated variance unique to depression symptoms. βc is the moderated variance components common to the moderator (i.e., standard of living or educational attainment) and depression symptoms. -2LL = negative 2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion (lower values reflect a better fit); Δ -2LL = likelihood ratio chi-square (X^2) test comparing the difference in -2LL of the models; Δ *df* = difference in degrees of freedom of the models; p = *p*-value of the X^2 test.

Table S7. Bivariate analyses between standard of living and depression symptoms, and between educational attainment and depression symptoms.

	Cross-Twin Cross-Trait Correlations	rA	rC	rE
Standard of Living				
MZM	-0.10 [-0.17, -0.02]	-0.15 [-0.28, -0.12]	1.00 [-1.00, 1.00]	-0.14 [-0.21, -0.06]
DZM	-0.08 [-0.17, -0.02]			
MZF	-0.03 [-0.10, 0.03]			
DZF	-0.05 [-0.13, 0.04]			
DZOS	-0.07 [-0.14, 0.01]			
Educational Attainment				
MZM	-0.09 [-0.16, -0.01]	-0.31 [-0.52, -0.14]	1.00 [-1.00, 1.00]	-0.01 [-0.07, 0.09]
DZM	-0.05 [-0.14, 0.05]			
MZF	-0.13 [-0.19, -0.08]			
DZF	-0.10 [-0.18, -0.02]			
DZOS	-0.02 [-0.09, 0.05]			

Note. MZM = monozygotic male; DZM = dizygotic male; MZF = monozygotic female; DZF = dizygotic female; DZOS = dizygotic opposite sex. rA = genetic correlation; rC = shared environment correlation; rE = non-shared environment correlation.

Supplementary 8: Code for biometric moderation models

Scripts for the biometric moderation models can be found here: https://github.com/isabellabadini/ses-depression-cotass

Appendix C

Supplementary materials for Chapter 4

Table S1. Estimated marginal means (EMMs) of internalising symptoms at different levels (-1 SD/mean/+1 SD) of education across different levels (-1 SD/mean/+1 SD) of continued participation.

	Educational attainment		
	-1	0	1
	EMM (95% CI)	EMM (95% CI)	EMM (95% CI)
Continued participation			
-1	0.19 (0.18, 0.21)	0.06 (0.05, 0.07)	-0.08 (-0.10, -0.06)
0	0.10 (0.09, 0.12)	-0.00 (-0.01, 0.00)	-0.11 (-0.13, -0.10)
1	0.01 (-0.01, 0.03)	-0.07 (-0.08, -0.05)	-0.15 (-0.17, -0.13)

Note. SD = Standard deviation; 95% CI = 95% confidence intervals

Table S2. Fit statistics for univariate ACE models for continued participation, educational attainment, and internalising symptoms.

		Model Fit					
Measure	Model	-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Continued participation	Full ACE	109141.4	37478	198151.4	-	-	-
	AE	**109141.4**	**37479**	**109149.4**	**-7.28e-10**	**1**	**1.00**
	E	109232.9	37480	109238.9	91.55	1	1.09e-21
Educational attainment	Full ACE	105178.5	36758	105188.5	-	-	-
	AE	**105178.5**	**36759**	**105186.5**	**0.06**	**1**	**0.80**
	E	105647.1	36760	105653.1	468.65	2	1.72e-102
Internalising symptoms	Full ACE	99441.99	34524	99451.99	-	-	-
	AE	**99441.99**	**34525**	**99449.99**	**-8.97e-08**	**1**	**1.00**
	E	99506.74	34526	99512.74	64.76	2	8.68e-15

Note. ACE model contains additive genetic (A), shared environmental (C), and nonshared environmental (E) parameter estimates. AE model contains AE parameter estimates; Best-fitting model in bold letters. -2LL = negative 2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion (lower values reflect a better fit); Δ -2LL = likelihood ratio chi-square (X^2) test comparing the difference in -2LL of the models; Δ *df* = difference in degrees of freedom of the models; p = *p*-value of the X^2 test.

Table S3. Fit statistics from the extended bivariate ACE models of the relationships between continued participation and educational attainment or internalising symptoms

		Model Fit					
Measure	Model	-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Continued participation – educational attainment	**Full ACE**	**213038.3**	**74233**	**213064.3**	**-**	**-**	**-**
	AE	213066.4	74234	213090.4	28.14	1	1.13e-07
	E	213136.9	74235	213158.9	98.65	2	3.79e-22
Continued participation – internalising symptoms	Full ACE	208318.80	71999	208344.8	-	-	-
	AE	**208318.80**	**72000**	**208342.8**	**0.01**	**1**	**0.92**
	E	208336.1	72001	208358.1	17.28	2	1.77e-04

Note. ACE model contains additive genetic (A), shared environmental (C), and nonshared environmental (E) parameter estimates. AE model contains AE parameter estimates; Best-fitting model in bold letters. -2LL = negative 2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion (lower values reflect a better fit); Δ -2LL = likelihood ratio chi-square (X^2) test comparing the difference in -2LL of the models; Δ *df* = difference in degrees of freedom of the models; p = *p*-value of the X^2 test.

Table S4. Fit statistics from the extended bivariate ACE models of the relationships between educational attainment and internalising symptoms in the full sample and subsample of individuals who participated at 8-years

		Model Fit					
Full sample (*N* = 2424)	Model	-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Education – internalising	Full ACE	203917.4	71279	203943.4	-	-	-
	AE	**203918.2**	**71280**	**203942.2**	**0.77**	**1**	**0.38**
	E	203946.3	71281	203968.3	28.92	2	5.26e-07
Subsample (*N* = 14046)	Model	-2LL	df	AIC	Δ -2LL	Δ *df*	*p*
Education – internalising	Full ACE	87071.68	31920	87097.68	-	-	-
	AE	**87073.64**	**31921**	**87097.64**	**1.96**	**1**	**0.16**
	E	87078.70	31922	87100.70	7.02	2	0.03

Note. ACE model contains additive genetic (A), shared environmental (C), and nonshared environmental (E) parameter estimates. AE model contains AE parameter estimates; Best-fitting model in bold letters. -2LL = negative 2 log likelihood; df = degrees of freedom; AIC = Akaike's Information Criterion (lower values reflect a better fit); Δ -2LL = likelihood ratio chi-square (X^2) test comparing the difference in -2LL of the models; Δ *df* = difference in degrees of freedom of the models; p = *p*-value of the X^2 test.