THE ENVIRONMENT AND CHILD DEVELOPMENT:
A MULTIVARIATE APPROACH

Amy Johnson

Hughes Hall

Cognition and Brain Sciences Unit

Medical Research Council

University of Cambridge

'Children are the world's most valuable resource

and its best hope for the future'

John F. Kennedy

ABSTRACT

The environment that a child grows up in has a profound effect on their child development. For example, key outcomes such as academic ability or behaviour, cognitive ability and the neurobiology of the brain have been found to be associated to a child's environment. However, the factors that make up a child's environment are highly complex and yet the majority of research treats SES as a single number. In addition, the environment is related to several aspects of a child's development, yet there is very little research considering how these multiple levels of development relate to each other and interact. This thesis builds on the current literature by investigating how multiple aspects of a child's environment combine to create an environmental profile that is associated with positive child development. We endeavour to address three questions:

1) Which environmental factors most strongly relate to a child's academic ability, behaviour, cognitive ability and neural development?

2) Does the wider environment mediate the relationship between standard measures of SES and child development?

3) How might the environment impact academic and behaviour outcomes? In particular, is this relationship mediated by a child's cognition or the structural and functional connectivity of their brain?

7-11 year-old children (N=97) and their caregivers took part in this study. Several environmental domains and child behaviour were assessed through caregiver and child questionnaires. Academic and cognitive ability were measured using behavioural assessments. Resting state functional connectivity was measured using a magnetoencephalography (MEG) scan and structural connectivity was measured in an optional MRI scan on a separate visit (N=87).

Partial Least Squares (PLS) methods identified significant relationships between the environment and child development. Multiple environmental domains were found to be reliably related to each aspect of child development. Furthermore, the wider environmental domains mediated the association between SES measures and each aspect of child development. Finally, cognition and the structural connectivity of a

child's brain mediated the association between the environment and academic outcomes. This was not found for the behavioural outcomes.

This thesis provides key advances towards addressing the considerable methodological challenge presented in the investigation of the complex relationships between a child's environment and multiple aspects of their development. We believe that this work will complement the available research to date and provide important detail to enable practitioners and policymakers to better support children at risk from disadvantaged environments.

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

SES	Socioeconomic status
MRI	Magnetic resonance imaging
AIC	Akaike information criterion
AWMA	Alloway Working Memory Assessment
BIC	Bayesian information criterion
BRIEF	Behaviour Rating Inventory of Executive Function
CCA	Canonical correlation analysis
DTI	Diffusion tensor imaging
EEG	Electroencephalography
FA	Fractional anisotropy
FAMD	Factor analysis of mixed data
fMRI	Functional magnetic resonance imaging
LASSO	Least absolute shrinkage and selection operator
MEG	Magnetoencephalography
MI	Multiple imputation
MICE	Multiple imputation via chained equations
PCR	Principal component regression
PLS	Partial least squares
RG	Resource Generator-UK
RGCCA	Regularized Generalized Canonical Correlation Analysis
ROI	Region of interest
SD	Standard deviation
SDQ	Strengths and Difficulties Questionnaire
sPLS	Sparse partial least squares
TBSS	Tract-based spatial statistics

LIST OF APPENDICES

1 INTRODUCTION

1.1 The environment and child development in Cognitive Neuroscience

With nearly 4 million children living below the poverty line, the UK has one of the highest rates of child poverty in the industrialised world (End Child Poverty, 2015). The environments that these children grow up in often lack access to the basic necessities such as food, energy and housing (Evans, 2004; UNICEF Office of Research, 2017). These environments are also characterized as having lower levels of social and cognitive stimulation, higher levels of stressful events, poorer quality childcare, schooling and accommodation and lower access to green spaces and safe places to play (Brito and Noble, 2014; Griggs and Walker, 2008).

Growing up under these circumstances can have a devastating and long-lasting effect on almost every area of a child's development. This impact begins before birth and accumulates across a child's lifetime. For example, poverty is associated with higher levels of infant mortality, premature births and lower birth weight. As a child develops, low SES is related to a host of health conditions including anaemia, diabetes, obesity, lead-poisoning, chronic asthma, poor dental health and number of accidents, hospital visits and days off school (Brooks-Gunn and Duncan, 1997a; Case and Paxson, 2006; Evans and Kim, 2007; Griggs and Walker, 2008).

A cognitive neuroscience approach provides evidence of this at multiple levels. For example, key outcomes such as academic ability or behaviour, cognitive ability and the

neurobiology of the brain have been found to be associated to a child's environment. This section will provide an overview of the current field across each of these levels of analysis, and its implications of our current understanding, before outlining the core aims of this thesis and how these will be delivered.

1.1.1 Academic outcomes

Children growing up in a low SES environment are consistently found to obtain poorer academic outcomes throughout childhood. This attainment gap can be seen from a very young age. For example, in a large study of school readiness in Scotland, children receiving free school meals (FSM) were 0.69 standard deviations (SD) behind in reading and 0.68 standard deviations behind in maths when they first begin school, in comparison to children that did not receive FSM (Scottish Government, 2005). In a longitudinal study following children from 12 to 36 months old even small gains in income produced significant gains in school readiness (Dearing et al., 2001).

In 2016 UK national statistics indicated that children in the first years of school who received FSM were on average 4.3 months behind their peers (Office for National Statistics, 2016a). This gap persists and widens at every Key Stage of schooling (KS), until, by the end of secondary school, children from disadvantaged backgrounds are 19 months behind, as can be seen in Figure 1 (Andrews et al., 2017). In short, children receiving FSM begin school at a disadvantage, they have a slower rate of progress throughout school, and this results in a cumulative effect over time. There are numerous good examples of this, taken from multiple different points within the education system. Secondary school students receiving FSM increasingly lag behind their peers by a rate of around 2 months each year (Office for National Statistics, 2016a). The pattern is repeated in higher education with pupils being 2 times more likely to progress to higher education courses if they are from non-disadvantaged backgrounds in comparison to those receiving FSM (Department for Education, 2017a). Of those that do progress, the fact that only 4.9% of students receiving FSM achieve 3 A's at A level in comparison to 11% of students not receiving FSM means that they are less likely to attend the more selective higher education providers (Department For Education, 2017). In addition, children from low SES backgrounds are also over 2 times more likely to have Special Educational Needs (SEN) (Shaw et al., 2016) and 4 times more likely to receive permanent or fixed term expulsion (Department for Education, 2017b).

The impact of poverty on educational attainment has lifelong implications. Educational skill and attainment is highly important for entry to work, job stability and progressions. Adults who experienced child poverty are more likely to be out of work or working in unskilled and more poorly paid jobs (Griggs and Walker, 2008). Not only are they more likely to be in poverty as an adult, their children will grow up in a low SES environment, repeating the cycle of poverty (Griggs and Walker, 2008; Smith and Middleton, 2007). As educational attainment is integral to mediating social mobility and enabling an individual to escape this cycle, closing the attainment gap remains one of the core aims of practitioners and policy makers.

Attainment gap in months between FSM and non FSM pupils

19.3	Key stage 4
9.5	Key stage 2
5.4	Key Stage 1
4.3	Early years
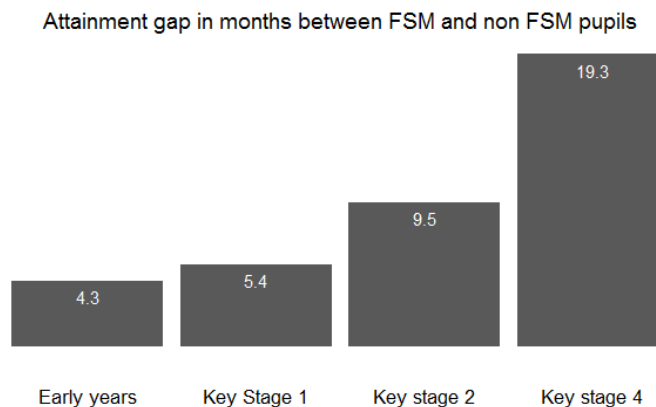
Figure 1 The attainment gap in months between children receiving free school meals and other pupils at each stage of a child's schooling.

1.1.2 Behavioural outcomes

Problematic emotional and behavioural outcomes, commonly considered in terms of externalising (e.g. hyper-activity, aggression and rule-breaking) and internalising (e.g. anxiety, social withdrawal, low self-esteem and depression) behaviours, are frequently associated with growing up in a deprived environment. For example, of the 55 papers that met the inclusion criteria of the systematic review by Reiss, 52 studies found a significant relationship between low SES and behaviour and mental health problems in children aged between 4-18. On average children from disadvantaged backgrounds were two-three times more likely to develop mental health problems. In another systematic review of 30 research reports in pre-school aged (3-5yrs) children, around 30% of children from low SES backgrounds were reported to have behavioural problems in comparison to between 6-8% of children in the general population. Higher incidences of behavioural problems associated with the environment are also reported in adolescents, as reviewed by (Dashiff et al., 2009).

The causal nature of this association has been investigated through randomised controlled trials, and other studies that exploit changes in income beyond the family's control such as changes to benefit systems. For example, Cooper et al. found 7 papers that recorded significant effects of income. A $1000 increase in income was associated with an effect size of 0.09-0.24 SD in measures of behaviour (Cooper and Stewart, 2013). Furthermore, this association is typically found to be stronger for externalising behaviours than internalising behaviours (Costello et al., 2003; Milligan and Stabile, 2011; Strohschein, 2005).

Note that the timing and duration of poverty have been found to moderate this association. Typically, the effects of poverty are stronger for children in early childhood in comparison to adolescence and for those that experience poverty for longer periods of time (Reiss, 2013). For example, Duncan et. al. found that children who were persistently poor were significantly more likely to display both externalising and internalising behavioural problems at age 5 than children who experienced short term poverty only in a longitudinal study (Duncan et al., 1994). The greater the frequency of poverty during a child's early life was significantly predictive of poorer externalising and internalising behaviour at age 14 (Najman et al., 2010a, 2010b). A longitudinal study found that persistent poverty in 4-8 year olds was strongly predictive of internalizing behaviours, even after controlling for the child's current poverty level. Externalising behaviours were only found to be related to current poverty (McLeod and Shanahan, 1993). Note that a similar study on an older age group of 8-11 year olds found the reverse: current poverty was associated with internalizing difficulties and persistent poverty was associated with externalising behaviour (Strohschein and Gauthier, 2017). In summary, low SES is associated with externalising and internalising behaviour problems, but the effect is moderated by participant age and timing of poverty.

1.1.3 Cognition

In addition to the impact of poverty on academic and behavioural outcomes, the negative impact of growing up in a deprived environment has also been found across assessments of cognitive skills. For example, infants as young as 6 months from a low socioeconomic (SES) background perform significantly worse than infants from a high SES background in the A not B task, an early predictor of executive functions (Lipina et al., 2005). This gap in executive functions appears to widen with age and low SES

children have significantly poorer inhibitory control (Farah et al., 2006; Lipina et al., 2013; Noble et al., 2005; Sarsour et al., 2011), cognitive flexibility (Czernochowski et al., 2008a; Kishiyama et al., 2009; Noble et al., 2005; Sarsour et al., 2011; Sheridan et al., 2012), executive planning (Lipina et al., 2013), attention (D'Angiulli et al., 2008; Kishiyama et al., 2009; Lipina et al., 2013; Mezzacappa, 2004; Neville et al., 2013; Stevens et al., 2009; Weatherholt et al., 2006), and working memory (Evans and Schamberg, 2009; Farah et al., 2006; Kishiyama et al., 2009; Lipina et al., 2013; Noble et al., 2007; Sarsour et al., 2011). By the age of 5, growing up in low SES circumstances is associated with a reduction in IQ of 6-13 points (Brito and Noble, 2014) and a wealth of evidence points to a large proportion of this being due to environmental rather than simply genetic effects alone (Capron and Duyme, 1989; Hamadani et al., 2014; Hanscombe et al., 2012; Marcus Jenkins et al., 2013; Turkheimer et al., 2003).

The greatest differences between low SES and high SES children are consistently found to be in measures of language (Farah et al., 2006; Noble et al., 2007). For example, children from low SES backgrounds have significantly poorer expressive and receptive language (Farah et al., 2006; Fernald et al., 2013; Jednorog et al., 2012; Kishiyama et al., 2009; Noble et al., 2007, 2005) and phonological awareness (Noble et al., 2007, 2006a, 2005; Raizada et al., 2008).

1.1.4 Neural measures

In recent years, a number of studies have adopted tools and methods from cognitive neuroscience to investigate the relationship between child development and the environment (see these references for a number of recent reviews: (Brito and Noble, 2014; Hackman et al., 2010; Hackman and Farah, 2009; Johnson et al., 2016; Lipina and Evers, 2017; Lipina and Posner, 2012; Raizada and Kishiyama, 2010; Ursache and Noble, 2016)). In general, these employ a univariate approach, in which the structure, task-related function or resting-state activity of brain areas are compared between low and high SES children. The following section outlines this research to date.

1.1.4.1 Brain structure

The structure of a child's grey and white matter is fundamental to their healthy development. Several associations between brain structure and SES have been identified in areas across the brain. Many of these studies have focussed on localized differences between low and high SES children. For example, particularly strong positive relationships between grey matter volume and SES have been identified in the

hippocampus (Butterworth et al., 2012a; Hair et al., 2015, 2015; Hanson et al., 2011a; Jednorog et al., 2012; Luby et al., 2013; Noble et al., 2015, 2012a; Staff et al., 2012) and amygdala (Butterworth et al., 2012a; Hair et al., 2015; Luby et al., 2013; Noble et al., 2012a). However, differences have also been found in widespread regions across the brain. For example, low SES is associated with thinner cortical surface areas and volume in the frontal (Hair et al., 2015; Hanson et al., 2012; Noble et al., 2015) (Jednorog et al., 2012; Lawson et al., 2013; Raizada et al., 2008) parietal (Hanson et al., 2013), temporal (Jednorog et al., 2012), insular (Jednorog et al., 2012) and anterior cingulate(Lawson et al., 2013) cortices. In a similar way, SES is associated with widespread white matter integrity in distributed networks across the brain (Dufford and Kim, 2017; Gullick et al., 2016; Ursache et al., 2016).

These SES-related structural differences have been related to key developmental outcomes in a small number of studies. For example, a recent cross-sectional study of 389 4-22 year olds found structural differences in brain areas associated with school readiness, including reduced grey matter volumes in the frontal and temporal cortices and the hippocampus in children from lower income families. These differences were found to account for between 15-20% of the income-related academic achievement gap seen in the sample (Hair et al., 2015). In a similar study of 1099 3-20 year olds by Noble. et. al., parental education and income were found to be logarithmically related to cortical surface area, with the relationship between poverty and brain structure strongest in children from the most disadvantaged backgrounds (Noble et al., 2015). The association was also found to be greatest in regions in the frontal, temporal and parietal cortices, and mediated the relationship between income and performance on inhibitory control and working memory tasks.

However, it is important to note that there are relatively few studies investigating associations between the environment and brain structure and several findings remain inconsistent. For example, results are mixed as to whether differences in global measures of brain structure such as grey and white matter volumes exist between low and high SES children (Butterworth et al., 2012a; Cavanagh et al., 2013; Jednorog et al., 2012; Krishnadas et al., 2013; Lange et al., 2010). Other studies have found either no differences or the inverse relationship between local grey matter volumes such as in the hippocampus (Butterworth et al., 2012a) and amygdala (Butterworth et al., 2012a; Hanson et al., 2011a; Noble et al., 2015, 2012b). In another example, Jednorog et al found local differences in grey matter volume in regions across the brain but no

differences in white matter properties (Jednorog et al., 2012). These inconsistencies are likely to be due, in part, to differences in methods, participant ages, SES measures and relatively small sample sizes across the studies.

1.1.4.2 Task-related brain activity

Task based neuroimaging is used to identify brain regions functionally involved in completing a specific task. Many brain areas have been associated with SES-related differences in patterns of brain activation during cognitive, academic and behavioural tasks (Johnson et al., 2016). For example, robust differences between low and high SES participants are frequently found in activation within and between the amygdala and prefrontal regions during tasks relating to emotion regulation and social exclusion (Gianaros et al., 2008; P. Kim et al., 2013; Muscatell et al., 2012; Taylor et al., 2006). In another example, fMRI during a learning task indicated that low SES children displayed significantly more activity in the right middle frontal gyrus in particular, in addition to the supplementary motor area, basal ganglia, bilateral inferior frontal gyrus, anterior cingulate cortex and significantly less activity in the superior frontal sulcus in comparison to high SES children (Sheridan et al., 2012). They hypothesise that this is driven by low SES children spending more time learning the rule associations than the high SES group and found that complexity of family language was related both to parental SES and pre-frontal activation. In two other studies investigating SES related differences during language tasks, low SES was related to lower asymmetry in inferior frontal gyrus during a rhyming task (Raizada et al., 2008) and lower left fusiform activity during phonological task in children with low phonological awareness (Noble et al., 2006b). Finally, whilst the majority of studies focus on fMRI imaging, differences in EEG measured brain activity has also been found. For example, in a series of EEG experiments, evidence was found for the relationship between SES and a participant's ability to selectively attend to stimuli based on the difference in ERP's between attended and non-attended stimuli. This was particularly focussed over frontal areas and present even in the absence of behavioural differences (D'Angiulli et al., 2012, 2008; Stevens et al., 2009)

1.1.4.3 Resting-state brain activity

In addition to structural neuroimaging and task-related functional neuroimaging, spontaneous and task-independent brain activity measured when the brain is at rest has also been an area of rapidly growing interest in recent years. A small number of studies

have investigated the relationship between SES and resting-state data using both EEG (Brito et al., 2016; Otero, 1997; Tomalski et al., 2013; Tomarken et al., 2004) and fMRI (Barch et al., 2016; Chan et al., 2018; Smith et al., 2015; Sripada et al., 2014). For example, SES-related differences have been found as young as 6-9 months with infants from low SES families showing lower levels of frontal gamma power, measured by EEG whilst watching a video (Tomalski et al., 2013). Differences in resting state EEG have been found to persist into childhood, particularly across frontal electrodes (Otero, 1997; Tomarken et al., 2004).

Recently, the relationship between SES and the correlation in brain activity between different regions, known as functional connectivity has been investigated. For example, Stripada et al. have found that childhood SES is significantly related to default mode network resting-state activity using fMRI in adults and that this in turn is associated with higher stress levels in a social stress test (Sripada et al., 2014). Another fMRI study of 7-11 year olds found that low SES was associated with reduced connectivity between the hippocampus and amygdala and several other parts of the brain including the superior frontal cortex, lingual gyrus, posterior cingulate, and putamen (Barch et al., 2016). However, there remains very few studies, the majority of which have focussed on a single network or connectivity with localized regions.

In summary, SES related differences in structure and function are found at a variety of locations across the brain. However, this field is in its infancy, and few studies have been conducted in comparison to other areas of child development. Little has been done to unify these findings or to consider how the different brain areas interact using multivariate approaches. This will be discussed further in section 1.2.

1.1.5 Childhood Interventions

Sadly, all too often, the impact of growing up in a deprived environment sets the trajectory for an individual's life, resulting in the perpetuation of poverty from generation to generation (Corak, 2006). However, this does not have to be the case. Childhood interventions can be an effective step in partially offsetting the negative impacts of poverty on child development, apparent both in terms of short-term effects and in providing resilience in the future, ultimately enabling people to break out of this cycle of poverty (Herrod, 2007; Howard-Jones, 2014). Of the wide variety of interventions available, many have been found to impact a child's development across

multiple levels including their academic and behavioural outcomes, cognition and, although there is only limited research, in some cases at a neural level.

As early childhood poverty in pre-school-aged children (around 3-5yrs) is often particularly devastating to a child's development (Brooks-Gunn and Duncan, 1997a; Cooper and Stewart, 2013), a great deal of attention is given in particular to pre-school interventions. This 'critical period' is likely to be partly due to the rapid brain development that occurs during this time and the fact that a pre-schoolers day-to-day life is dominated by the home environment. In a systematic review of 123 quasi-experimental and experimental studies of US pre-school interventions based in educational centres, average effect sizes were small: 0.14SD in academic progress, 0.16SD in socio-emotional development and 0.23SD in cognition were found between age 3 to adulthood (Camilli et al., 2010). Larger effect sizes were found in a systematic review of 56 studies on 30 interventions in mostly low income countries that only included randomized control trials and strong quasi-experimental design trials (Nores and Barnett, 2010). In this, medium effect sizes of 0.41SD for school progress, 0.27SD for behaviour, 0.35SD for cognition and 0.23SD for health were found for taking part in early educational interventions.

In primary and secondary schools, the introduction of the Pupil Premium in 2011 by the UK government has provided additional resources to schools to help tackle the attainment gap through childhood interventions. The Education Endowment Foundation (EEF) Teaching and Learning Toolkit provides an excellent summary of ongoing research into international evidence for key educational interventions, to help schools identify potential interventions (Education Endowment Foundation, 2018). Out of the 34 intervention styles investigated, 28 indicate positive effects. The most effective interventions provide an average improvement of 8 additional months' progress over the course of a year, such as high-quality feedback, meta-cognition and self-regulation interventions. The next most effective include homework, peer-tutoring, mastery learning, reading comprehension strategies, collaborative learning and oral language interventions, each providing an average of 5 months additional development over the course of one year. Note that each of these interventions is described by the EEF as being low cost, calculated at under £70 per year per student.

It is important to keep in mind that the positive effects of an intervention can be seen not only for the individual, in terms of behaviour or educational attainment, when delivered on a scale they have the potential to deliver strong economic benefits for the

country. Poverty in the UK is estimated to cost the nation £78 billion pounds: higher poverty levels resulting in increased spending in public services and the knock-on effect of lower earnings and employment prospects (Bramley et al., 2016). It's further estimated that £1 in in every £5 spent on public services are aimed at making up for the damage that poverty inflicts on a person's life (Bramley et al., 2016). There is therefore a strong argument that, if more were spent on childhood interventions that were effective in reducing the damage that poverty inflicts, then the savings in public services and the knock-on effects of poverty could far outweigh the cost of the interventions. As an example, a benefit-cost analysis of the Abecedarian pre-school program was conducted using a randomised control study that followed individuals until age 21. When benefits such as the increase in earnings of the participant and future generations, additional maternal employment during pre-school years, reduced school education costs, healthcare and adult welfare costs was offset against project costs (and the additional cost of higher education), the benefits of outweighed the costs by 2.5:1 in the Abecedarian project and an enormous 17:1 in the Perry Pre-school Programme (Barnett and Masse, 2007a, 2007b). Furthermore, it would be hoped that as these generations go on to parent the following generations, there would be a reduced need for interventions in the future, providing a strong incentive for governments to devote resources to childhood interventions in the present.

The question of which childhood interventions are the most effective (and least expensive) remains an area of great importance in research in the 21st Century. Whilst many demonstrate positive results, typically only relatively small to medium effect sizes have been found. Others, including large scale, highly expensive programs such as the Head Start and Early Head Start pre-school programs in the US show little or no positive impact under randomized control trials (Puma et al., 2010; The Early Head Start Research Consortium, 2002). In particular, there is a great need to better determine the relationship between the environment and child development, in order to highlight the key environmental factors of risk and resilience that have the greatest impact. Identifying these key environmental factors has the potential to help practitioners identify children who may be most at risk of poor child development, select the interventions that may be most effective for them as individuals and ultimately enable researchers, policy makers and practitioners to design more effective interventions by identifying targets for intervention. Cognitive neuroscience is an integral area of

research in addressing these questions and therefore has a vital role to play in narrowing the gaps in development between children in different socioeconomic circumstances.

1.2 Factors to address

Given that various key environmental factors can impact upon child development, this thesis will seek to address two key issues typically seen in the field of cognitive neuroscience. Firstly, the factors that make up a child's environment are highly complex and yet the majority of research treats SES as a single number. Secondly, the environment exerts its impact at multiple levels of a child's development but there is a lack of research investigating its impact across multiple levels, including at the neural level, or considering how different levels may interact.

1.2.1 Multiple environmental factors

The majority of cognitive neuroscience research that has explored childhood SES has either treated it as a single measure, such as parental education, occupation or income or as a composite measure of these. A frequent criticism of the field is that this approach is too reductive (Raizada & Kishiyama, 2010). In reality, a child's environment is governed by a variety of factors that interact and typical measures of SES do not capture this effect on a child's development. For example, these might include aspects of cognitive stimulation, health, parenting attitudes and styles, subjective SES, neighbourhood SES, relationships and support and childcare (Ackerman and Brown, 2006a; Duncan et al., 2017; Farah, 2017; Lipina, 2016; Ursache and Noble, 2016). Many of these environmental factors are not typically captured by a single measure such as parental income. Indeed, some of these measures may exert a different or stronger impact on a child's development than standard SES measures (D'Angiulli et al., 2012). We might also expect the SES measures to exert an effect on child development via driving differences in variables in the wider environment. Investigating the mediating effect of environmental variables on the relationship between parent income, education and occupation and child development will provide evidence for potential mechanisms that underlie the relationships. Furthermore, many of the wider environmental factors have the potential to be easier and less expensive to alter through childhood intervention than, say, increasing a family's income (Education Endowment Foundation, 2018). By better identifying the impact of multiple aspects of a child's environment and how these combine to create an environmental profile that is conducive to positive child

development we hope to provide an evidence-base for developing effective interventions for a child's development.

A related issue is the lack of studies that consider how the environmental profile is reflected in individual differences. There is a tendency in the research to treat children from a similar SES band as a single group, and simply compare childhood outcomes between different SES groups. However, we know that there is a great deal of variability in the development of children from a similar SES background, with some children more at risk of poor developmental outcomes and others more resilient to these risks (Werner, 1993).

This is apparent in our pilot data, conducted with 310 children (158 males (51%), mean age 9:7±1:2 and age range 6:11-11:9 (years:months)) from 3 Cambridgeshire primary schools. Children completed 4 timed paper and pencil tasks in the classroom. Reading Fluency and Maths Fluency from the Woodcock Johnson III Form C Tests of Achievement (McGrew et al., 2007; Woodcock et al., 2001) battery were used to obtain age standardised measures of academic ability. The Series and Classifications subtests from the Cattell Culture Fair scale 2 (Cattell, 1949, 1963) were used to measure fluid IQ. Note that these could not be age standardised as we could not administer the full battery due to time constraints, but there was no relationship between FSM status and age. When grouped by FSM status, children receiving FSM (low SES) performed significantly worse than children not receiving FSM (high SES) in each subtest. However, the density plots in Figure 2 demonstrate that there is considerable variability within each group and overlap between the groups. For example, a proportion of children from the low SES group achieve some of the highest scores, demonstrating a remarkable resilience to the negative effects of low SES. We hypothesize that factors from the wider environment that are not captured well by simplistic measures of SES may be driving this individual variability, causing some children from the same SES group to be more at risk, and others to show resilience to these risks. Identifying the environmental profiles that drive these individual differences is key to identifying individual risk factors and designing interventions that may foster resilience to these specific risks.
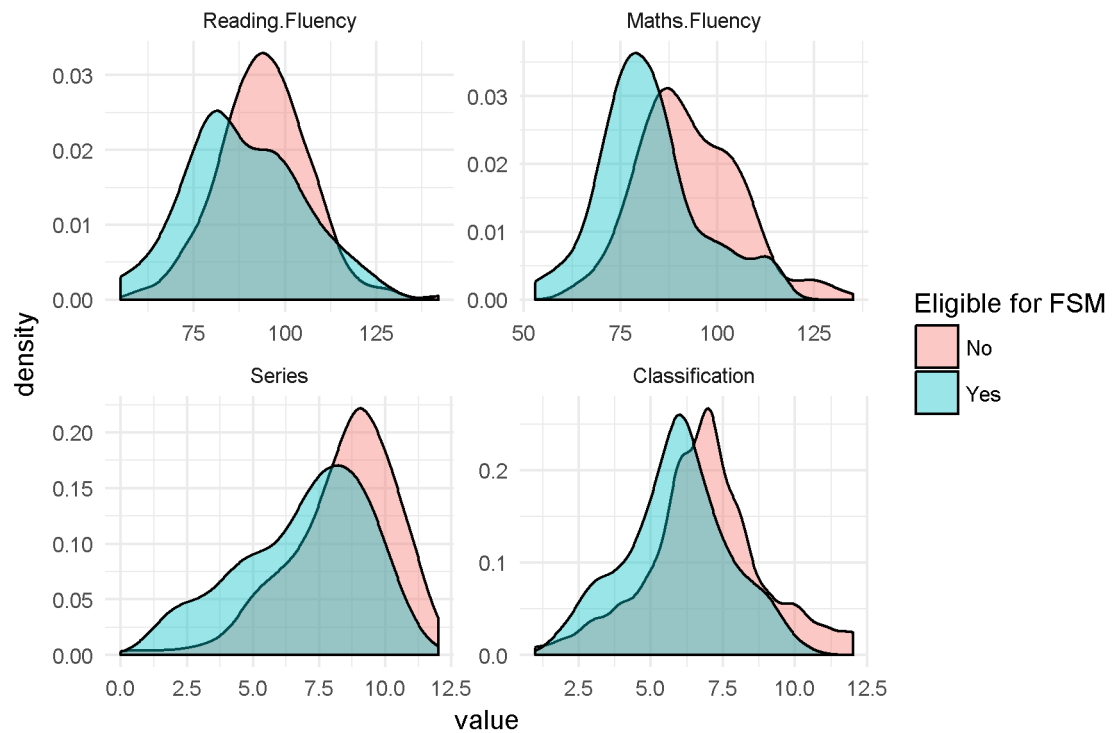
Figure 2 Density plots of reading, maths and IQ scores in our pilot dataset, grouped by FSM status.

1.2.2 Multiple levels of child development

As is evident from the literature reviewed in sections 1.1.1-1.1.4, the environment is closely related to several aspects of a child's development, including their academic ability, behavioural outcomes, cognition and brain structure and function. However, there is very little research considering how these different aspects relate to each other and interact. To the best of our knowledge there are no datasets that combine robust measures of a child's academic and behavioural outcomes, cognition and neural factors, or explore how each of these interact to create a broad picture of how a child's development is influenced by various aspects of their environment. In measuring at each of these levels of analysis we aim to identify whether and how different environmental influences impact at different levels of child development. Additionally, by considering whether the impact of the environment on key outcomes such as academic ability and behaviour is mediated by a child's cognition or neural measures, we hope to shed light on the potential mechanisms through which the environment may exert its effect.

In particular, studies that include neuroimaging methodologies are rare. In their recent review of the field Raizada and Kishiyama noted that:

"There is a remarkable disconnect at present between the large amount of behavioural data which is available and the almost complete absence of corresponding neural data. This presents quite a research opportunity" (Raizada and Kishiyama, 2010).

Given that academic and behavioural outcomes depend on the development of a child's brain (Pugh et al., 2001; Tau and Peterson, 2010), neural measures hold particular potential for gaining insight into the underlying mechanisms by which the environment might exert an effect on these childhood outcomes. For example, it is likely that a positive environment will support structural or functional developments in a child's brain which, in turn, facilitate learning. In addition, academic and behavioural outcomes reflect the combined function of multiple underlying cognitive and socio-emotional systems (Hackman et al., 2010). It is possible that the environment might be related to a specific underlying system or have a more global association across the brain. Indeed, for the same child outcome, there may be a variety of different mechanistic pathways by which the environment may exert an effect that are not apparent in the outcome measures alone. Neuroimaging helps us to tease these systems apart by identifying the brain regions and pathways most strongly associated with the environment. Furthermore, 'specific' effects in the brain often translate to multiple behavioural outcomes (Richmond et al., 2016). Identifying the neural factors that are related to the environment, may provide insight into why such a wide variety of childhood outcomes are all impacted by the environment. Additionally, the scores achieved in outcome tasks are dependent on the cognitive strategies employed by participants and the specific task used, making interpretation of individual differences difficult. In contrast, structural and resting-state functional neuroimaging allows us to probe the intrinsic brain architecture underlying child development, independent of task or strategy. Finally, it is also possible that neural measures may be more sensitive to the effects of the environment, or perhaps relate to specific factors in the environment that are not apparent in behavioural performance (Raizada and Kishiyama, 2010).

Furthermore, in the few papers that have employed neuroimaging methods as described in section 1.1.4, the majority employ univariate approaches in which signals from the brain are analysed voxel by voxel (Raizada and Kishiyama, 2010). However, it is likely that, particularly when considering multiple environmental factors, the relationship between the environment and brain development will not be isolated to particular localised brain areas (Raizada and Kishiyama, 2010; Richmond et al., 2016). Methods in neuroscience have been rapidly advancing over recent years, enabling us to better

investigate whole-brain organisation and the networks that are formed by different areas of the brain acting in concert, and yet very little research has capitalised on these advances in the field of child development and the environment. This thesis aims to apply some of the latest advances in neuroimaging such as whole-brain structural and functional connectomics to provide a detailed analysis of how the environment relates to child brain development and how this might translate into the widespread impact of the environment on academic and behavioural outcomes.

1.2.3 Challenges to overcome

There are key challenges that have resulted in these factors remaining largely under researched to date. These will need to be addressed in order to identify the impact of multiple environmental factors on a wide range of developmental measures. Firstly, a number of difficulties arise from high dimensionality in a dataset. In particular, many standard statistical methods such as linear regression cannot be applied in circumstances in which the number of variables exceeds the number of observations. This is a common issue in neuroscientific research due to the prohibitive expense and time-consuming nature of neuroimaging, often preventing studies from having a large number of participants.

The number of variables can be reduced by applying dimension reduction techniques such as averaging or factor analyses. However, dimension reductions techniques tend to be performed on a single dataset, such as the environmental factors without considering the relationship to other datasets in the analysis such as the academic, behavioural, cognitive or neural measures. Many of these techniques are variance based and work by creating a smaller number of dimensions that best explain the variance in the original dataset. However, it may be that the variables that maximise the explained data in one set are not the same variables that best explain the relationship between two different sets of data. Again, this is likely to result in weaker relationships being found between datasets such as the environmental factors and the developmental measures, despite the existence of a strong relationship between individual variables between these datasets. In addition, any specific effects such as a particularly strong relationship between the environment and one aspect of cognition or a specific brain area may be reduced or washed-out by averaging over several measures in a dataset such as across the whole brain or all cognitive measures.

Furthermore, dimension reduction has the potential to reduce the interpretability of results. For example, in the single child neuroimaging study that we know of in which a wide range of environmental factors were considered, Jenkins et. al. applied a principal components analysis to 14 SES and 28 family stress indicators to identify three dimensions that best explained the child's environment before investigating how these related to IQ, grey and white matter in children (Marcus Jenkins et al., 2013). Whilst some of the resulting components consisted only of questions related to a specific aspect of the environment, others, including the first and only component to show a significant mediation effect, were made up of a mix of relatively dissimilar aspects of the environment, making interpretation and application to interventions challenging.

Secondly, there are likely to be relatively strong relationships within a dataset, such as between the environmental factors or within the neuroimaging measures. This multicollinearity not only creates challenges in identifying which environmental factors are the most important predictors of child development, but also has a detrimental effect on many statistical analyses such as regression methods, which rely on the assumption that all predictor variables are orthogonal, resulting in models being unpredictable and unstable (Gorman, 2010; Gunst and Mason, 1979). For example, in a re-analysis of the relative importance of different SES factors in a linguistic study, Gormen demonstrated that multicollinearity between the SES measures had resulted in biased results in the original study (Gorman, 2010). A common solution to this issue is either to remove variables until only an independent set of predictors remain, which raises difficulties in how to select variables or to perform dimension reduction techniques that typically suffer from the issues described above. Therefore, in aiming to identify the most important factors of risk and resilience in children from a set of related environmental factors, a great deal of care must be taken in how to analyse this data in the presence of multicollinearity.

Finally, in order to better understand how the environment relates to several aspects of a child's development, it would be very useful to be able to model the relationships between multiple (more than two) datasets. This, for example, would allow us to investigate whether there are cognitive or neural factors that mediate the relationship between childhood environment and key child outcomes such as academic attainment or behaviour. However, the issues of dimensionality and multicollinearity also prevent the use of simple mediation analyses and covariance based structural equation models typically used to investigate relationships between several sets of variables on the same

Amy Johnson – September 2018

individuals. As a result, the few studies to date that have investigated mediating effects between the environment and multiple aspects of child development have been limited to either reducing a dataset to a single measure or a very small number of measures (Hair et al., 2015; Hanson et al., 2012; Holz et al., 2015; Krishnadas et al., 2013; Lipina et al., 2013; Luby et al., 2013; Noble et al., 2015).

1.3 Statistical Methods

The challenges of high dimensionality, multicollinearity and complex, multi-dataset studies have largely resulted in research that considers only a limited set of environmental factors and aspects of child development. However, these issues are not unique to this field and, particularly in light of the 'big data explosion' occurring across the board, are increasingly being encountered and addressed in almost every area of research. As a result, our ability to analyse data under these circumstances has been rapidly progressing and a number of methodological developments hold potential in helping us better understand how the environment shapes child development. Two areas that show particular promise and growing popularity in addressing these challenges are the fields of partial least squares (PLS) (Krishnan et al., 2011) and feature selection, outlined in the sections below. These two areas offer complementary techniques; PLS for characterizing the environmental profiles most strongly related to child development and LASSO for selecting a smaller set of environmental factors within this profile that are the most useful for predicting child development.

1.3.1 Partial Least Squares

PLS was developed by Herman Wold in the 1970s to identify the sets of variables that best model the relationship between two or more datasets (Wold, 1982, 1975). In contrast to typical regression methods, this relationship is modelled via an iterative method that both allows the number of predictor variables in a dataset to be greater than the number of observations and makes it possible to find stable and interpretable solutions when variables within a dataset are collinear (Wegelin 2000; Yeniay and Goktas 2002).

In addition, PLS is a dimension reduction technique that identifies a smaller set of latent factors for each dataset, made up of a linear combination of the variables in the dataset. However, in contrast to typical dimension reduction techniques such as Principal Component Regression (PCR), which seeks to maximise the variance explained within

each individual dataset, PLS seeks to maximise the covariance explained between two or more datasets. As such, the goal of PLS is to identify a smaller set of latent variables that best explain the relationship between datasets. The degree to which a variable in a given dataset loads onto a latent variable can be interpreted as the relative importance of that variable to describing the relationship. For example, Smith et al. recently applied Two-block Mode B PLS, more commonly known as Canonical Correlation Analysis (CCA), to data from healthy adults in the Human Connectome Project in order to investigate the relationship between 280 behavioural and environmental variables and participants' functional connectomes (Smith et al., 2015). They used this to demonstrate that one specific pattern of brain activity was significantly related to a set of environmental and behavioural factors and used the loading of each of these onto latent variables to identify which were most related to individual differences in the functional connectome. In another recent example, Ziegler et al. used PLS to identify the cognitive profiles in children that most strongly covary with regional grey matter volumes (Ziegler et al., 2013). Primarily, they found that the strongest relationship was between parent's ratings of behaviour related to executive function and grey matter volumes in widespread regions across the brain. Thus, in the context of this study, PLS is a technique that will allow us to identify the profiles of factors in a child's environment and aspects of their development that are most strongly related.

Finally, PLS can be extended to simultaneously model relationships between more than two sets of data, such as in PLS path modelling (Tenenhaus et al., 2005; Wold, 1975) or Regularised General Canonical Correlation Analysis (RGCCA), the generalised model that includes PLS techniques (Tenenhaus and Tenenhaus, 2014, 2011) in the context of high dimensional, multicollinear data. Many of the developments in this field are very recent but the techniques are rapidly gaining popularity in various fields using big, multi-table datasets including genetics (Garali et al., 2017; Günther et al., 2014; Meng et al., 2016; Rohart et al., 2017; Tenenhaus and Tenenhaus, 2014), chemical spectroscopy (Smilde et al., 2005) and business and market research (Hair et al., 2011; Henseler et al., 2016). However, to our knowledge, this will be the first study that aims to apply PLS methods to investigate the relationship between multiple aspects of the environment and child development in the field of cognitive neuroscience.

1.3.2 Feature Selection

Whilst the PLS methods characterize the relative importance of each variable to the outcomes, it is also useful to consider which features are the most important to measure to best predict the markers of child development. Collecting a large amount of data is generally impractical, particularly for practitioners. It is also likely in the context of high dimensional, multicollinear datasets that the measurement of some variables may be redundant as the information is captured well by a smaller set of variables. Selecting a limited set of environmental variables that best capture the environmental variance most strongly related to child development will help practitioners and policy-makers to identify the key environmental factors to measure. For example, this might enable practitioners to identify which children might be most at risk of poor child development and might benefit most from early childhood interventions, based on their environment.

In the context of the current big data explosion, feature selection methods are rapidly being developed to deal with the challenges of high dimensional datasets (Bolón-Canedo et al., 2015). One of the most prominent and widely used amongst these is the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1994). This is a type of linear regression that applies L1 regularisation to shrink some of the regression coefficients towards zero, reducing the original set of predictor variables to create a simple and sparse structure. The variables amongst a highly correlated set that best capture the shared variance are selected, whilst redundant variables are removed from the regression whose variance is better explained by the selected variables. It is demonstrated to outperform many classic methods such as stepwise and backwards selection. Thus, LASSO feature selection provides a powerful tool for identifying a limited set of environmental factors that are the most important to measure to best predict the markers of child development in this study.

1.3.3 Causation vs Correlation

PLS methods and LASSO feature selection are valuable tools for identifying the key factors of a child's environment that most strongly relate to their development. However, it is important to note that these methods do not demonstrate causality between the environment factors and child development. Instead, they describe patterns of correlation between the environment and child development datasets. For example, a strong relationship does not imply that differences in an environmental factor causes differences in child development. Whilst we might hypothesize that this is the case, a

strong association may also be the result of a third factor that causes both the environment and development measures, reverse causation in which differences in child development cause differences in environmental measures or even coincidental correlation. There can be no conclusion drawn on whether a causative relationship exists or it's direction using PLS and LASSO selection.

The gold standard for establishing causality is to conduct randomised control trials. In this, groups of participants are matched on as many factors as possible whilst the variable of interest is varied between the groups. A subsequent difference between the groups demonstrates an effect. In addition, longitudinal studies in which the same individuals are followed over time can be used to provide evidence for causation as causes generally precede their effects. However, these techniques generally require a great investment of time, money and energy. It would be unfeasible to conduct experiments in this way using the many varied measures of a child's environment and development suggested in this thesis. Before such studies can be undertaken, the specific aspects of a child's environment that relate most strongly to their development must first be pinpointed. Thus, this cross-sectional study aims to provide a foundation for further studies into the cause and effect of the child's environment and their development by providing a valuable evidence base for future research to draw on. Characterizing how a wide range of different factors in a child's environment relate to markers of child development and identifying a smaller set of the most predictive environmental variables will identify key potential targets for further research.

1.4 The study

This thesis will investigate the relationship between a child's environment and their development in a sample of 7-11-year-old children (N=97). Multiple environmental factors will be assessed through parent and child questionnaires. Academic ability, child behaviour, cognition and functional and structural neuroimaging will be measured to reflect child development at multiple levels. Advances in the field of neuroscience and statistical techniques will be used to identify the set of factors that combine to create an environmental profile that relates to these aspects of child development. In particular, we endeavour to address three main questions:

1. Which environmental factors most strongly relate to a child's academic ability, behaviour, cognitive ability and neural development?

2. Does the wider environment mediate the relationship between standard measures of SES and child development?

3. How might the environment impact academic and behaviour outcomes? In particular, is this relationship mediated by a child's cognition or the structural and functional connectivity of their brain?

This thesis provides key advances towards addressing the considerable methodological challenge presented in the investigation of the complex relationships between a child's environment and multiple aspects of their development. We believe that this work will complement the available research to date and provide important detail to enable practitioners and policymakers to better support children at risk from disadvantaged environments.

2 METHODS

Children were invited to attend a 3 hour session at the Medical Research Council Cognition and Brain Sciences Unit, Cambridge, during which they took part in academic and cognitive assessments and a 1 hour magnetoencephalography (MEG) scan. During this, their primary caregiver completed questionnaires about their child's behaviour and a wide range of aspects of their home environment. In addition, families were also invited to attend an optional 1 hour structural magnetic resonance imaging (MRI) scan on a separate visit. In this chapter data collection is described, followed by the methods used to pre-process the data, including correcting for missing values and the factor analyses used to characterise the environmental domains and behavioural measures. Finally, the PLS and feature selection methods used in each analysis of the data in the following chapters are outlined.

2.1 Data collection

2.1.1 Participants

97 children took part in the main behavioural and MEG session (demographics given in Table 1). 87 children also completed the optional MRI scan on a separate visit. Families were recruited via local schools and advertisements in public places across the city. Recruitment of participants to scientific studies often suffers from a bias towards higher SES families. Therefore, in order to obtain a sample reflective of the spread of SES in the UK population, care was taken to reduce this bias as much as possible. Flyers inviting children to take part were sent through three schools that have over 27% of students receiving FSM. Flyers and posters were sent to all Sure Start Children's Centres in Cambridge City and South Cambridgeshire and to houses, GP's and library's in areas across Cambridge City that have the highest percentage of children living under the poverty line. Each family was reimbursed for their time (around £74 if one child took part, £112 for 2 children or £153 for 3 children for each family completing the full study). Parents provided written informed consent and this study was approved by the Psychology Research Ethics Committee at the University of Cambridge (Reference: 2015.11).

Table 1 Main sample demographics

Demographic	Main sample	Optional MRI
N total	97	87
n Males	45 (46%)	43 (49%)
n Right Handed	81 (84%)	71 (82%)
n Ethnicity: White-British Other Not Given	 82 (85%) 14 (14%) 1 (1%)	 73 (84%) 13 (15%) 1 (1%)
Age (years:months)	Mean = 9:10 +/- 1:5, Range = 6:11-12:9	Mean = 9:11 +/- 1:6, Range = 6:11-12:9

2.1.2 Academic outcomes

Several subtests from the Woodcock Johnson III Form B Tests of Achievement (McGrew et al., 2007; Woodcock et al., 2001) battery were used to assess age standardised academic ability. Each task took around 5 minutes to complete and the total academic and cognitive task session lasted 1.5 hour including breaks.

Reading skill was measured using the Letter-Word Identification, Passage Comprehension and Reading Fluency subtests. In Letter-Word Identification (untimed) participants read single words of increasing difficulty to assess their word identification skills. In Passage Comprehension (untimed) participants supply missing words in passages of increasing complexity to assess understanding of written text. In Reading Fluency participants read as many simple sentences as they can in 3 minutes and answer true/false questions to assess their ability to read simple sentences quickly.

Maths skill was assessed using the Calculation and Maths Fluency subtests. In calculation (untimed), participants complete a number of calculations of increasing difficulty. These are mostly numerical operations but participant's progress to geometric, trigonometric, logarithmic and calculus operations if appropriate. In Maths Fluency, participants answer as many addition, subtraction and multiplication

calculations as they can in 3 minutes to assess their ability to solve simple calculations quickly.

The subtests were combined to calculate a Reading and Maths score respectively using the age standardisation from the Normative Update (McGrew et al., 2007). In addition, they completed the Spelling subtest which can be combined with the scores from the Letter-Word Identification and Calculation subtests to obtain an 'Academic Skills' score, but this was not used in the analyses presented here, because we opted to focus on literacy and numeracy skills.

2.1.3 Behaviour outcomes

The Behaviour Rating Inventory of Executive Function (BRIEF) (Baron, 2000) parent form for ages 5-18 was used to measure problems in child behaviour relating to executive functions. This is an 86 item questionnaire that can be divided into eight clinical subscales: Inhibit (assessing inhibitory control) , Shift (difficulties in shifting between different situations or activities), Emotional Control (difficulties in controlling emotional responses), Initiate (difficulties in initiating problem solving or activities), Working Memory (difficulties in sustained working memory), Plan/Organize (difficulties in planning or organising problem solving approaches), Organization of Materials (difficulties in organising their environment and materials) and Monitor (difficulties in monitoring their own behaviour). The 25 item Strengths and Difficulties questionnaire (SDQ) (Goodman, 1997) for 4-17 year olds was used to assess externalising (e.g. hyper-activity, aggression and rule-breaking) and internalising (e.g. anxiety, social withdrawal, low self-esteem and depression) behavioural difficulties. The 25 items are divided into 5 subscales: Hyperactivity (such as restlessness and attention), Emotional Symptoms (such as anxiety or depression), Conduct Problems (such as lying or temper), Peer Problems (such as bullying or lack of friends) and Prosocial Behaviour (such as being considerate and generous). The primary caregiver completed these questionnaires during the child's behavioural testing session.

2.1.4 Cognitive tasks

Several age standardised measures of cognitive ability were obtained using a battery of computerised, verbal and paper-and-pencil tasks. Each task took around 5-10 minutes to complete. Verbal and Fluid IQ were measured using the verbal and Matrix Reasoning subtests from the Wechsler Abbreviated Scale of Intelligence II (WASI-II) (Wechsler,

2011). Short-term and working memory (STM and WM) was assessed using four subtests of the Automated Working Memory Assessment (AWMA): Digit Recall, Backwards Digit Recall, Dot Matrix and Mr X (Alloway, 2007). These are designed to tap verbal STM, verbal WM, visuo-spatial STM and visuo-spatial WM respectively. Phonological processing was assessed using the Naming Speed subtest from the Phonological Assessment Battery (PhAB) (Frederickson and Reason, 1997).

2.1.5 Environmental factors

Several additional parental questionnaires were created to investigate multiple aspects of SES, including standardised measures of parental income, education, occupation and family subjective SES and other factors in the child's environment. These can be found in Appendix A.

2.1.5.1 Income

Net household equivalised income was used as it is the standard measure of income used by the UK government. This adjusts household income to account for the fact that households of different size and composition require different incomes to attain the same living standard, allowing comparison between households. Available yearly income (the net household income after tax, including any money received from benefits and child support payments if the parents of the child are separated and a payment arrangement is in place) is divided by the total household scale factor calculated from the Modified OECD equivalence scale (Anyaegbu, 2010) given in Table 2.

Table 2 Modified OECD equivalence scale. The factor for each individual in the household is summed and the net household income after tax is divided by this total.

Individual	Modified OECD equivalence scale
First adult	0.67
Subsequent adults or children aged 14+	0.33
Child aged 0-13	0.20

2.1.5.2 Caregiver education

Caregiver education for the primary and secondary (where applicable) caregivers were scored on a 7 point scale shown in Table 3 based on the education scale from the Hollingshead four-factor index of SES (Hollingshead, 1975) and adapted for the British school system. Average household education was found for families of two caregivers to allow comparisons with single parent families.

Table 3 Highest level of caregiver education

Highest level of Education	Score
Primary school	1
Some secondary school	2
CSEs/O levels/GCSEs/Level 1 or 2 NVQ	3
AS or A levels/ Level 3 or 4 NVQ/ Level 3 BTEC/International Baccalaureate	4
Certificate of higher education/diploma/foundation degree	5
Bachelor's degree/ Professional certificate	6
Master's Degree/PGCE /PhD/Doctorate	7

2.1.5.3 Caregiver occupation

Historically, caregiver occupation has typically been based on subjective measures of prestige such as the Duncan Socioeconomic Index (SEI) and revised versions (Duncan, 1961; Stevens and Featherman, 1981), one of the most commonly used measures of SES available. However, in addition to being highly subjective, prestige is continually changing for the vast number of occupations available and it has been argued that such methods rapidly become obsolete (Featherman et al., 1975; Hauser and Warren, 1997). Instead, this study used the National Statistics Socio-economic Classification (NS-SEC) (Office for National Statistics, 2016b; Rose and Pevalin, 2003), the official socio-economic classification used in by the UK government, based 'objectively' on employment relations and hierarchy. The three class hierarchical scale was used to class

each caregivers occupation using the Standard Occupational Classification (SOC) 2010 scores (E. L. and S. A. Office for National Statistics, 2016) into: Higher managerial, administrative and professional occupations, Intermediate occupations and Routine and manual occupations. The average score for the household (either single parent or two caregivers) was used.

2.1.5.4 Subjective SES

Subjective SES aims to obtain a measure of how a participant subjectively views their position in the 'social hierarchy'. Whilst there is typically a high correlation between this and objective SES measures such as income, occupation and education, subjective SES has been found to predict psychological functioning and health-related factors over and above these standard SES measures (Adler et al., 2000; Yanagisawa et al., 2012). The MacArthur Scale of Subjective Social Status (Adler et al., 2000) was used in this study to measure subjective SES. Participants are given a picture of a 10 step ladder and asked:

'Think of a ladder with 10 steps representing where people stand in the UK (changed from 'United States' in original scale). At step 10 are people who are the best off – those who have the most money, the most education, and the most respected jobs. At step 1 are the people who are worst off – those who have the least money, least education, and the least respected jobs or no job. Where would you place yourself on this ladder?'

2.1.5.5 Social Resources

The Resource Generator-UK questionnaire (Webber and Huxley, 2007) was used to measure primary caregiver's access to social resources within their social network. This includes 27 yes/no questions about the skills and resources that a participant might have access to in their social network within one week. These questions are further divided into 5 subscales: domestic resources, expert advice, personal skills and problem-solving resources. There are an additional 13 questions to assess the participants own skills.

2.1.5.6 Neighbourhood SES

The 2015 English Indices of Deprivation were used to obtain measures of relative deprivation of the Lower layer Super Output Area (LSOA) that each participant is currently living in, based on their postcode (Office for National Statistics, 2015). The 37 separate indicators are combined into 7 domains: income, employment, education, health, crime, barriers to housing and services and living environment.

2.1.5.7 Holmes Rahe Stress Scale

Parental stress was measured using the Holmes Rahe Life Stress Inventory, a 43 question scale of key stressful life events (Holmes and Rahe, 1967). Primary caregivers were asked to complete this for both the past year and the duration of their child's life.

2.1.5.8 Additional measures: custom questionnaires

In order to obtain additional measures of a child's wider environment, a number of questionnaires were created. These questions were designed to measure factors in the following 12 domains: parent health, child health, attitudes to child's education, use of other languages, attitudes to reading, time spent with family and friends, subjective opinions of the quality of family relationships, childcare, child use of technology, discipline, rules and chores and parental attitude to others in their neighbourhood. Children also completed a questionnaire designed to measure how they feel about their family, friends, education and self and these questions were divided amongst the relevant domains. These questionnaires included questions adapted from questions in the Growing up in Australia questionnaires (Australian Institute of Family Studies, 2017) and the Understanding Society questionnaires (Institute for Social and Economic Research, 2015). The questionnaires can be found in Appendix A.

2.1.6 MEG Scan

MEG data were acquired using a high-density VectorView MEG system (Elekta-Neuromag) that contains a magnetometer and two orthogonal planar gradiometers at each of 102 locations. Five head position indicator (HPI) coils are attached to the child's head in order to monitor the child's head movements throughout the recording. Their relative position is recorded using a 3D digitizer (FASTRACK, Polhemus) in addition to over 50 additional points distributed over the scalp. Pulse was measured using an electrode attached to each wrist and eye movements were recorded using horizontal and vertical electrooculograms. Children were monitored by video camera throughout the scan. Children completed a resting-state scan, a WM dot matrix task and a phonological word learning task. Only the resting-state scan was analysed within this thesis, and the task data will be used in future analyses.

During the 9 minute resting-state scan, children were asked to sit as still as possible, close their eyes and let their mind wander. This was followed by the first block of the WM task, two blocks of the phonological oddball task and finally, the second block of

the WM task. Each of these blocks lasted about 10 minutes and short breaks were taken in-between each block.

The WM was adapted from the task used by Barnes et al. (Barnes et al., 2016) and has been shown to be a sensitive measure of individual behavioural and neural differences in WM ability based on standardised assessments from the AWMA. In the WM task children had to remember the spatial location of 4 target dots presented sequentially in a matrix. After the four matrices have been presented a final matrix appears with a question mark placed at one of the previous target locations. The participant has to recall the order that the targets were presented in and respond using a button box according to when in the sequence the target appeared in the probed location. In addition, in half of the trials two distracting dots of a different colour are also presented. The child is instructed to ignore these distracting dots and to focus on the target dot. It has previously been found that performance is reduced in the presence of distracting dots, and it is thought that this provides a measure of our ability to selectively attend to tasks (McNab et al., 2015). Each block contains 100 no distractor trials and 100 distractor trials. The experimental paradigm can be seen in Figure 3.
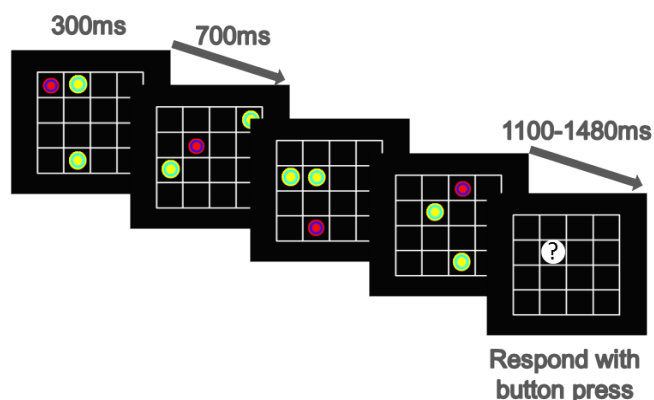
Figure 3 A single distractor trial from the WM task. The no-distractor trials contain only the red targets.

During the phonological oddball task participants are presented with three types of auditory stimuli binaurally through headphones: a novel pseudo-word standard stimuli ('boak'), a known word deviant ('boat') and a novel pseudo-word deviant ('boap') in a ratio of 6:1:1. After beginning with a train of 10 standard stimuli to habituate participants to the standard stimuli, there are either 2, 3, 4, or 5 repetitions of the standard stimuli in between 150 trails of each type of deviant, randomly interspersed. The interstimulus interval (from the end of one stimulus to the onset of the next) is kept

constant at 800ms. Hawkins et al. have kindly let us use the stimuli used in their similar oddball experiment (Hawkins et al., 2015). The stimuli were created from naturally spoken words by cross-splicing the final voiceless stop consonants from each word (the /t/, /p/ and /k/) onto the end of the first consonant-vowel token ('boa') from the known word 'boat'. This ensures that each stimulus is identical up until the start of the final phoneme at 370ms after onset so that we can identify the exact point at which the different stimuli can be distinguishable. Participants were instructed to ignore the sounds whilst watching an engaging silent video. The task was broken into two blocks of 12 minutes each. The experimental paradigm can be seen in Figure 4.

$$S\ S\ D_W\ S\ S\ S\ D_P\ S\ S\ D_P\ S\ S\ S\ S\ D_W\ S$$

Figure 4 An example from the multi-feature oddball paradigm. Known word deviants (D_W) and novel pseudo-word deviants (D_P) are randomly interspersed between strings of pseudo-word standard stimuli (S).

2.1.7 MRI Structural Scan

After being familiarised with the MRI process in the on-site mock scanner, children took part in a 20 minute MRI scan whilst watching a DVD. T1-weighted and diffusion-weighted images were acquired using a during a 20 min Siemens 3T Tim Trio system (Siemens Healthcare, Erlangen, Germany), using a 32-channel quadrature head coil. A Magnetisation Prepared Rapid Acquisition Gradient Echo (MP RAGE) sequence with 1mm isometric image resolution, 2.98ms echo time and 2250ms repetition time was used to acquire whole-brain T1-weighted volume scans. An Echo-planar diffusion-weighted sequence with an isotropic set of 60 non-collinear directions, using a weighting factor of $b = 1000s/mm^{-2}$, interleaved with 4 T2-weighted ($b = 0$) volumes, 60 contiguous axial slices, isometric image resolution of 2mm, 90ms echo time and 8400ms repetition time was used to acquire the diffusion weighted images.

2.2 Pre-processing of the questionnaire and behavioural task data

2.2.1 Missing data

Whilst the proportion of missing data in each questionnaire and behavioural item was relatively low, the majority of items contained some missing data (average percentage of missing data = 4%, range = 0-30%). As a result, careful treatment of these missing values is required to ensure that subsequent analyses are unbiased. Methods have moved

far beyond using either listwise or pairwise deletion or simply replacing missing values with the mean, median or mode of a variable which are known to bias subsequent results. Amongst the alternatives that exist, multiple imputation (MI) via chained equations (MICE) (van Buuren, 2007) remains one of the most accurate, flexible and widely used methods for dealing with missing data (Azur et al., 2011). In MI, the statistical uncertainty in imputations is accounted for by estimated missing values m times to create a number of imputed datasets. This results in m complete datasets, where the missing values have been drawn from a distribution and so vary between imputed datasets. Subsequent analyses are performed on each of these datasets in turn and key results of interest are pooled across datasets using Rubin's rules (Rubin, 2004) as follows:

Suppose we are estimating a scalar of interest, Q, such as a factor loading. The ith (i=1,2…m) imputed dataset gives an estimate, Q_i with associated variance U_i. The pooled estimate is simply the average over N datasets:

$$Q = \frac{1}{m}\left(\sum_{i=1}^{m} Q_i\right)$$

Note that wherever the results of interest might be arbitrarily rotated from imputation to imputation, such as the factor loadings in a factor analysis, orthogonal Procrustes rotation is applied to account for this before pooling the results. This is discussed in more detail in sections 2.2.3 and 2.3.

In order to obtain a valid estimate of the standard error, the between-imputation variation has to be included in addition to the within-imputation variation, in order to account for the uncertainty in the imputation model. The within-imputation is the average of the imputed dataset variance estimates:

$$W = \frac{1}{m}\left(\sum_{i=1}^{m} U_i\right)$$

The between- imputation variance is calculated as:

$$B = \frac{1}{m-1}\sum_{i=1}^{m}(Q_i - Q)^2$$

These are combined to calculate the total variance as:

$$T = W + \left(1 + \frac{1}{m}\right)B$$

Note that Rubin's Rules require the scalar of interest, Q, to follow a normal distribution or t-distribution. However, this is not the case for p values, which are uniformly distributed between 0 and 1 under the null hypothesis. It would be useful to combine the set of m p values resulting from m analyses on m imputed datasets to obtain an overall estimate of the statistical significance of a result. In 2010, Licht proposed the 'z-transformation procedure' (Finch, 2016; Licht, 2010) based on Rubin's ideas as a solution to this issue as follows:

The p values are transformed to have a normal distribution using

$$Z_m = \Phi^{-1}\left(1 - p_m\right)$$

where p_m is the p value for imputed dataset m and Φ^{-1} is the inverse normal distribution. The mean z value, $\bar{z}$ and between imputation variance, B, are calculated using Rubin's Rules above. The within imputation variance, W, is 1 by definition as the p values have been transformed to a standard normal distribution. The total variance is then calculated as

$$T = 1 + (1 + 1/m)B$$

Finally, the pooled p value is calculated by comparing with a normal distribution of mean 0 and variance T:

$$p_{pooled} = Prob(z(0,T) \geq \bar{z})$$

MICE has emerged as one of the primary methods of MI. In this, all missing values are initialized with a random value. Next, an imputation model for the first variable with missing data is fitted using all other variables and the missing values in this variable are replaced with this imputation. This is repeated in turn for each variable, using the current updated imputation values for all other variables and iterated until the imputed values converge. Finally, the whole process is repeated m times to produce m imputed datasets. Because each variable is imputed using its own model, MICE is particularly effective for mixed data types, such as in our dataset which contains a mixture of continuous, polytomous and dichotomous items (Azur et al., 2011; van Buuren, 2007). Several methods are available for building these models, each more or less applicable to particular situations. MICE using a Random Forest algorithm was selected for our dataset as it has been found to be the most effective method when there may be non-

linear relationships between variables, when there are mixed data types and when there are more variables than observations (Stekhoven and Bühlmann, 2012).

MI using Random Forest MICE was implemented using the mice package in R (van Buuren and Groothuis-Oudshoorn, 2011). All variables that will be used in the subsequent analyses, including dependent variables, must be included in the imputation to avoid underestimating the strength of relationships in statistical analyses (Moons et al., 2006) and so we ran the imputation on all 168 questionnaire and behavioural items with each type set to continuous, ordinal or factor. In addition, it is also considered good practice to include auxiliary variables that might improve the imputation estimates and so we included gender, age and handedness (Azur et al., 2011). Although MI has been found to be an effective method even at large proportions of missing data, only a few items contained a large amount of missing data and so the two items with over 15% of data missing (BMI and rules about which games a child can play on a console) were removed from the analysis in addition to the three participants that had over 15% of data missing (Figure 5). The resulting dataset was imputed 5 times.
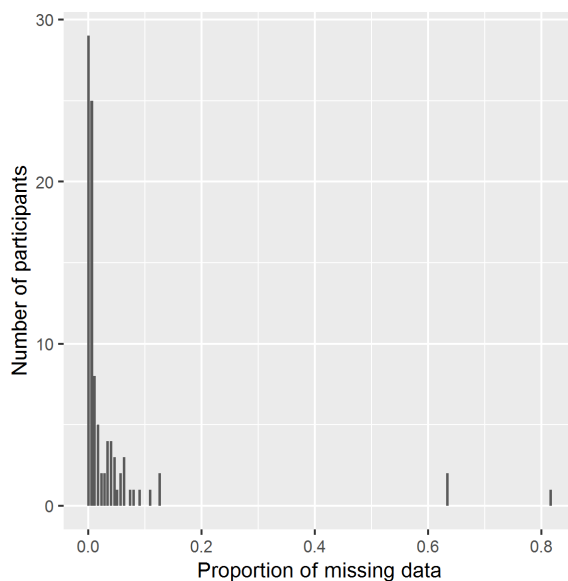
Figure 5 Distribution of missing data for participants. The three participants with over 15% of their data missing were removed from further analyses

2.2.2 Removing skewed items and extreme outliers

Of the 168 environment and behavioural items used in the imputation, a further 19 were excluded due to having a severely imbalanced distribution, measured as having >90% of subjects all having the same value for this item. These were the same across each

imputation in addition to the original dataset (listed in Appendix B, Table 8). A test for extreme outliers, as used in (Smith et al., 2015), was applied. Namely, if the maximum squared difference between the individual participant values and the median value for an item is greater than 100 times the mean value then the item is considered to have extreme outliers. None of the remaining 149 items in the original dataset or each of the imputed datasets met this condition.

2.2.3 Characterising the environment

2.2.3.1 Factor analysis

A child's environment can be thought of as being made up of different distinct but related domains such as their health, their parent's health and wellbeing and attitudes to things like education or discipline. These cannot be observed and measured directly and can be thought of as unobservable latent factors, F. We can, however, ask questions relating to each domain, such as asking about the number of hours sleep the child gets each night. The answer given, X_i, can be thought of as being influenced in part, by a set of general health factors, F, in addition to some other factors unique to the question, U_i, and can be expressed as a combination of these:

$$X_i = Fa_i^T + U_i$$

A factor analysis can then be used to estimate the factor loadings, a_i, and from this we can estimate the factor score for each child and so obtain a measure of their environment in the given domain, such as health. For example, Figure 6 demonstrates how 13 different items from the questionnaires in this study load onto a factor, producing a score that describes a child's general health level.
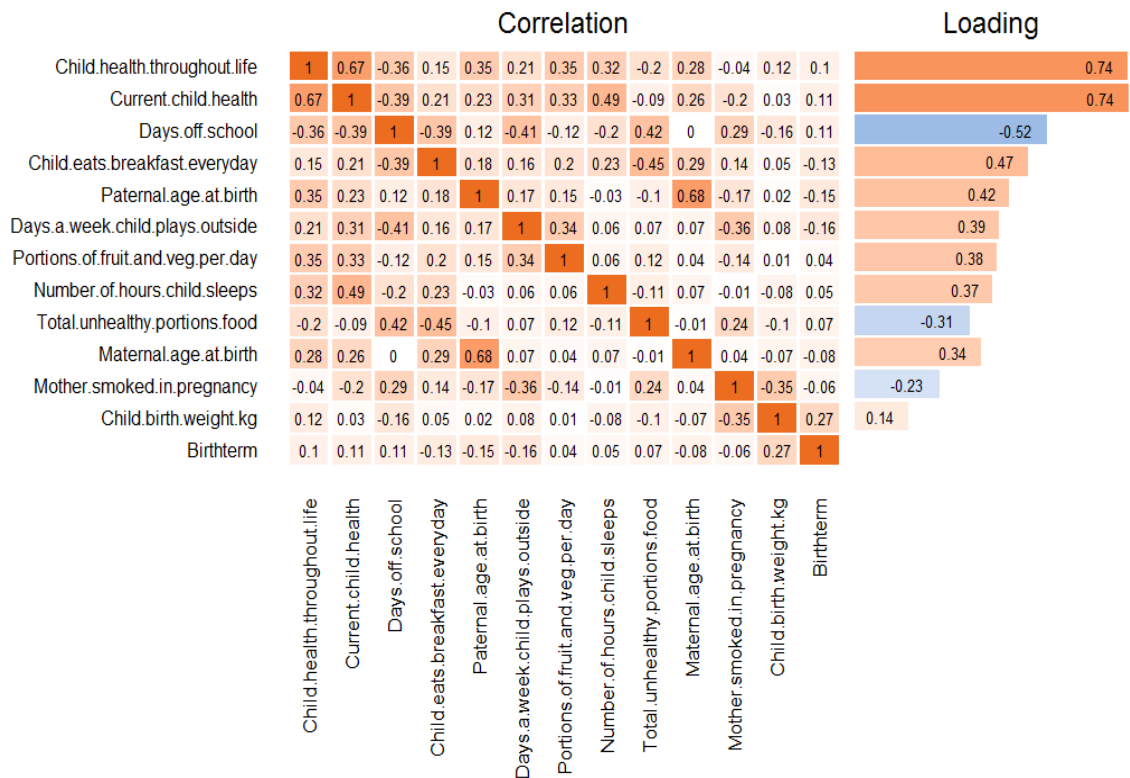
Figure 6 Example factor analysis for the child health domain. Correlation and factor loading for the multiple questionnaire items that relate to child health

Application of this method to our questionnaire data has a number of advantages. Firstly, combining multiple questions of the same underlying construct to characterize an environmental domain provides a more reliable measure than using individual questions alone. Secondly, in identifying a smaller number of underlying factors of a child's environment, the results are more manageable and interpretable than if we simply used the individual questions. Finally, it aids application of this research by practitioners who are unlikely to have the exact questions or behavioural measures obtained but may have other data relating to the domains.

2.2.3.2 Factor analysis of mixed data

The data is made up of a mix of dichotomous yes/no questions, polytomous Likert-type questions and continuous questions. A factor analysis of mixed data (FAMD) can be used to analyse all of these data types within one factor analysis (Pages, 2004). In this, the Pearson correlation matrix that is used in an ordinary factor analysis is replaced with a mixed correlation matrix in which Pearson correlations are calculated between continuous variables, polychoric between polytomous, tetrachoric between dichotomous and polyserial and biserial between the continuous variables and the other types. The

factor analysis can then be conducted on this correlation matrix. The questionnaire items were divided into the domains and the mixed correlation matrix for each group of questions in a domain was calculated using the 'mixed.cor' function from the 'psych' R package. Principal Axis Factoring was applied to each of the imputed datasets for a given domain.

2.2.3.3 Factor analysis and multiple imputation

There is rotational ambiguity in the factor analysis solution due to the fact that if the factor loadings and scores are rotated by a rotation matrix T, the model fit remains the same (Ginkel and Kroonenberg, 2014). This can be demonstrated using the fact that orthogonal rotation matrices have the properties $TT^{-1} = I$ and $T^{-1} = T^T$. Let F and a_i be rotated by T: $T^{-1} = T^T X_i = aTT^{-1}F + U_i$

The imputed datasets are each slightly different due to variation in the imputed missing values, and so the factor models for each may be rotated, reflected and reordered (the factor components extracted may be in a different order, particularly if two explain similar amounts of variance) with respect to each other due to this rotational ambiguity. To account for possible rotations between the imputations, the factor loadings for each of the imputations were rotated using Orthogonal Procrustes rotation onto the loadings for the first imputation and the factor scores were recalculated for each (Ginkel and Kroonenberg, 2014; Goodall, 1991; Voillet et al., 2016). Orthogonal Procrustes rotation finds the rotation matrix that minimises the sum of squares distance between the loadings in one imputation and the corresponding loadings from another. This is then used to rotate the loadings to ensure that the factor space modelled is equivalent for each imputed set of factors.

2.2.3.4 Creating the environment domains

One method for selecting domains of a child's environment to consider might be to conduct one overall factor analysis and allow the data to identify combinations that maximise the explained variance. This is the method used by the only paper that we know of that investigated multiple SES measures in a similar way (Marcus Jenkins et al., 2013). However, it is also of utmost importance that the factors are as simple, interpretable and as useful to practitioners as possible. The complex factors that result when we feed all question items into one big factor analysis may explain the variance well, but are messy, with questions that relate to very different aspects of a child's environment loading onto one factor, making them uninterpretable.

Therefore, we balanced data-driven factors that maximise the variance explained, and theory-driven factors that cleanly describe an area of interest. Our quasi-data-driven approach first divided the questions into pre-determined key domains of interest (attitude to child education, attitude to neighbourhood, child health, discipline, rules and chores, neighbourhood SES, other language use and reading, parent health and wellbeing, social resources, family relationships and use of technology) before running a separate factor analysis on each of these domains. The domains and corresponding questionnaire items can be found in Appendix B, Table 9.

It may be that the variance captured by the individual items in a domain is not explained well by the first factor alone. However, retaining multiple factors for one domain will reduce the interpretability of the factor scores, particularly as some question items may load onto more than one factor. We took two theory-driven approaches in these cases. Firstly, if the factor structure under Promax rotation was clearly divided into sets of question items that made up meaningful sub-domains, a subsequent factor analyses was run on each sub-set of items separately and the domains were renamed. For example, the factor analysis of the 'other language use and reading' domain indicated that the items were best captured by two factors that neatly divided items into reading (first factor) and other language use (second factor) sub-domains. A factor analysis was run on each of these sets of items independently to create two separate domains; 'other language use' and 'reading at home'. Secondly, the items that either had very low loadings onto the first factor or whose relationship to the domain was unclear were excluded if their removal simplified the factor structure.

This process was repeated until each domain was explained well by one factor. A number of criteria can be used to indicate the best number of factors to extract. In this analysis, we primarily used the eigenvalue criteria, extracting factors whose eigenvalue was greater than 1. However, as this approach has been criticized for underestimating the number of factors to be extracted, we also used inspection of the scree plot and parallel analysis, which compares the eigenvalues obtained to those that would be found for a random dataset of the same size, to guide us. Where these secondary conditions were not met (e.g. there was no discernible elbow in the scree plot), we checked that the additional variance explained by two factors in comparison to the first factor was minimal.

A detailed description of the application of this process to the questionnaire items can be found in Appendix B, Section B.1. The resulting questionnaire items used for each

domain are given in the final column of Table 9, in in Appendix B. In addition, the factor loadings and the correlations between each item for each domain, averaged across imputations, can be seen in Figure 72 to Figure 85 in Appendix B, section B.2. The factor loadings for each domain were then be used to calculate a factor score for each child in each domain. This process resulted in a set of 14 scores on a continuous scale for each child that describe the main environmental influences that they are experiencing. The explained variance for each of the final factors ranged from 12% to 86% (mean= 38%) and can be found in Table 4. Note that this is likely to be relatively low given that the individual question items are noisy measures, often having a large amount of variance unique to each item.

Table 4 Percentage of variance explained by the factor extracted for each environmental domain , average over imputations

Domain	Proportion Variance Explained
Acquaintance skills (RG)	0.47
Attitude child education	0.12
Attitude to neighbourhood	0.46
Child health	0.19
Childcare	0.43
Discipline	0.34
Neighbourhood SES	0.58
Opinion family relationships	0.32
Other language use	0.86
Primary caregiver wellbeing	0.44
Reading at home	0.19
Rules and chores	0.33
Technology use	0.29
Time with family and friends	0.26

In addition, 7 questionnaire items were kept as individual measures. The key standard measures of SES: income, parent education, parent occupation and subjective SES, although highly correlated, were used separately as they represent distinct resources that may relate differently to the neural, cognitive and outcome data (Duncan and Magnuson, 2012). The number of caregivers, number of siblings and the average number of hours the primary caregiver works each week were also used as individual measures as they did not fit any of the domains, but we believe they are important measures to consider. Although some of these measures (parent education and occupation) are not continuous, they are hierarchical, with higher values relating to higher education level or occupation and so will be treated as continuous in subsequent analyses. Therefore, in total, 21 scores relating to different domains of the environment were calculated for each child.

2.2.4 Additional factor scores: child behaviour and cognition

Additional factor analyses were also run on the BRIEF subscales, the SDQ subscales, and the age-standardised AWMA Working Memory subtest scores to obtain data-driven scores for each of these areas for each of the imputed datasets. Each of the academic, behaviour and cognitive subscales are given in Appendix C, Table 10. The factor analyses indicated that a single factor was underlying each measure of child development, based on the criteria that only the first factor had an eigenvalue greater than 1, the elbow of the scree plot and the parallel analysis of random data. These factors were used to calculate a single BRIEF, SDQ and working memory score for each child. The BRIEF and SDQ scores were multiplied by -1 so that higher scores related to fewer behavioural difficulties. The factor loadings and the correlations between each subtest, averaged across imputations, can be seen in Figure 86 to Figure 88 in Appendix C, section C.1. The explained variance for each of the final factors can be found in Table 5.

Thus, each child had an additional set of 8 scores relating to two academic measures (Maths ability and Reading ability), two behaviour outcomes (BRIEF and SDQ) and four cognitive abilities (working memory, Matrix Reasoning (performance IQ), Vocabulary (verbal IQ) and Phonological processing speed).

Table 5 Percentage of variance explained by the factor extracted for each measure of child development , averaged over imputations

Domain	Proportion Variance Explained
BRIEF	0.60
SDQ	0.37
Working memory	0.35

2.3 Partial Least Squares

2.3.1 Partial least squares general methods

Partial least squared (PLS) methods were used to investigate the relationship between the environment, cognition, brain measures and the academic and behaviour outcome datasets. PLS acts to find a set of orthogonal latent variables for each dataset that best explain the covariance between datasets. Each latent variable is, in a similar way to PCA, made up of a weighted linear combination of the variables in the corresponding dataset:

$$L_j = X_j a_j$$

Where L_j is the set of latent variables for dataset X_j and a_j is the corresponding set of outer weights.

There are a number of methods available, based on Wold's original PLS algorithm that are applicable in specific situations depending on whether two datasets (PLS) or multiple dataset (e.g. PLS Path Modelling) are modelled, whether the goal is describing the relationships between datasets (e.g. Canonical PLS) or predicting one dataset from another (e.g. PLS Regression) and whether the outer weights are reflective (Mode A) or formative (Mode B, commonly known as Canonical Correlation Analysis).

In this study we used the Regularised Generalised Canonical Correlation Analysis (RGCCA) framework which is a recent generalisation that encompasses all of the above PLS methods in addition to several other related methods. This will be used to both model relationships between the environmental dataset and either single or multiple aspects of the child development. An important motivation for using RGCCA is due to a small adjustment to Wold's algorithm (specifically when in Mode A) which has long

been criticised due to the lack of analytical proof of convergence and the fact that it does not optimise a specific property. In RGCCA, through a slight adjustment to the normalisation of outer weights called 'new Mode A', it can be demonstrated that the algorithm is monotonically convergent, and the latent variables are selected to maximise:

$$max \sum_{j,k=1,j \neq k}^{J} c_{jk} g\left(cov(L_j, L_k)\right)$$

$$constrained\ by: (1 - \tau_j)var(L_j) + \tau_j \|a_j\|^2 = 1\ for\ j = 1,2 \dots J$$

Where J is the total number of blocks. By selecting these three parameters, c_{jk}, $g(x)$ and τ_j, RGCCA equates to a variety of different multi-block methods. c_{jk}, is the JxJ symmetric design matrix describing the network of connections between blocks, usually 1 between two blocks that are connected and 0 if they are not. $g(x)$ is the scheme function, chosen to be Horst's scheme $g(x) = x$ in this study as it results in maximising the covariance between each pair of latent variables across the different datasets. Finally, τ_j, defines the normalisation of the weight vectors, chosen to be $\tau_j = 1$ for (new mode A) in this study because the latent variables are considered to be an underlying factor that causes the measured variables rather than a formative model in which the measured variables cause the latent variables (mode B, $\tau_j = 0$) . Thus, this study produces results that are scaled but equivalent to both PLS and PLS Path Modelling with Horst's Function and new Mode A and maximises:

$$max \sum_{j,k=1,j \neq k}^{J} c_{jk} cov(L_j, L_k)\ constrained\ by: \|a_j\|^2 = 1\ for\ j = 1,2 \dots J$$

In other words, the algorithm finds sets of latent variables for each dataset that maximally explain the covariance shared between the datasets in the model. This solution is found via an iterative algorithm in which the latent variables are alternately estimated via the outer model (linear combination of weighted variables in block) L_j , and the inner model (based on the other latent variables connected to a given block in the design matrix) Z_j until convergence is achieved as follows.

Step A: Initialization

Choose arbitrary initial outer weights $a_j^{(0)}$ and normalize according to $\left\|a_j^{(0)}\right\|^2 = 1$. In this study we initialized using the SVD of each matrix with itself, as is common practice.

Repeat Steps B to D with s = 1, 2, 3 …

for blocks j = 1, 2, 3… J do

Step B: Compute the outer approximation

$$L_j^{(s)} = \pm X_j a_j^{(s)}$$

Step C: Compute the inner approximation

$$Z_j^{(s)} = \sum_{k<j}^{J} c_{jk} L_k^{(s+1)} + \sum_{k>j}^{J} c_{jk} L_k^{(s)}$$

Step D: Update the outer weights

$$a_j^{(s+1)} = \frac{X_j^T Z_j^{(s)}}{\left\|X_j^T Z_j^{(s)}\right\|}$$

end for

Until convergence of the outer weights

Step E: calculate the final set of latent variables

$$L_j^{(S)} = x_j a_j^{(S)}$$

If additional components are to be extracted, the variance explained in each data matrix, X_j, is removed by deflation. In this study canonical deflation is used as we wish to explain the shared variance between the datasets rather than predict one from another. With this, each X_j is predicted via regressing the dataset onto the corresponding latent variable and subtracting this estimate from X_j. Let 'r' be the component number:

$$X_j^{(r\ est.)} = L_j^{(r)} \left(L_j^{(r)^T} L_j^{(r)}\right)^{-1} L_j^{(r)^T} X_j^{(r)}$$

$$X_j^{(r+1)} = X_j^{(r)} - X_j^{(r\ est.)}$$

Steps A to E are repeated using $X_j^{(r+1)}$ for components r = 1, 2, 3... maximum number of components. The sets of latent variables extracted with each round are ordered by covariance explained so that the first component latent variables explain the highest

amount of covariance. Note that in most cases the maximum number of components is limited to the number of variables in the smallest dataset block.

The outer weights

An important property of RGCCA under new mode A is that the outer weights relate to the sum covariance between a given variable in a dataset and each of the latent variables for the other datasets. This is demonstrated below:

$$a_j = \frac{x_j^T z_j}{\left\| x_j^T z_j \right\|}$$

$a_j \propto X_j^T Z_j$ and using the fact that $X_j^T Z_j \propto cov(X_j, Z_j)$ this becomes

$a_j \propto cov(X_j, Z_j)$

As the inner approximation Z_j is equal to the sum of the latent variables, $L_{k \neq j}$, of the other datasets connected to X_j, this demonstrates that the outer weights indicate the relative importance of each variable in a dataset for explaining the relationship with the other datasets. This is particularly useful in the case of just two datasets as it means that a given outer weight is proportional to the covariance between the corresponding variable and the latent variable in the other dataset, $a_j \propto cov(X_j, L_k)$. This means that we can use the outer weights to identify the variables in the environment and child development measures that most strongly relate to each other.

Note that for second and higher components, the variance explained by the previous components has been removed. Consequently, the outer weights are proportional to the remaining covariance between the specified dataset and the other latent variables. Whilst it is interesting to identify these higher dimensions of covariance between datasets, this can mean that the importance of variables in the higher components is difficult to interpret and so interpretation of the key environmental variables that relate most strongly to the different aspects of child development will primarily come from the first component.

2.3.2 PLS and multiple imputation

For each analysis, the PLS model was found for each imputed dataset. Note that the sign of the latent variables are ambiguous $L_j = \pm X_j a_j$ which can result in the direction of the latent variables and outer weights being flipped between imputations. It is therefore important to correct for this before finding the average latent variables and outer

weights across imputations. In order to account for this rotation, we have used Procrustes rotation to find the sign that minimises the squared Euclidean distance between the outer weights of each imputation dataset and the outer weights of the first imputation. This is the case only when each vector of outer weights from one component at a time is rotated, as this reduces the Procrustes rotation matrix to a single number (either 1 or -1). After the signs ambiguity has been accounted for, results can be combined across imputations according to Rubin's rules (Bastien, 2008). To our knowledge, this is the first occasion in which PLS has been applied to a multiply imputed dataset. Functions from the MixOmics Package (Rohart et al., 2017) were adapted for the analysis of multiply imputed data using in-house scripts developed for this project in R which can be found at https://github.com/ajohnson62.

2.3.3 Assessing significance and reliability in Two-block PLS

When the relationship between two datasets is modelled, we can assess the significance of each set of latent variables for a given component by comparing the covariance between each pair of latent variables, $Cov\left(L_X^{(orig)}, L_Y^{(orig)}\right)$ to the covariance found for when there is no relationship between the datasets (Krishnan et al., 2011). This was tested using permutation testing in which the rows of one dataset were shuffled, PLS regression was refitted to this, and the covariance between the latent variables recalculated. This was repeated 1000 times to build up a null distribution of $Cov\left(L_X^{(perm)}, L_Y^{(perm)}\right)$ for each component. The set of latent variables for a component are considered to significantly explain the shared variance between the datasets if they have a significantly higher covariance than would be found under permutation of the data according to:

$$p = \frac{\left(Cov\left(L_X^{(perm)}, L_Y^{(perm)}\right) > Cov\left(L_X^{(orig)}, L_Y^{(orig)}\right)\right) + 1}{number\ permutations + 1}$$

p was calculated for each imputed set of data and combined using the Licht- Rubin method of applying a z-transform to each p, calculating the average and transforming back to account for the fact that p values under the null hypothesis will be uniformly distributed rather than a normal distribution (Licht, 2010).

We also want to be able to assess the reliability of the outer weights for each significant component. As is commonly done, we used a bootstrap analysis to estimate the standard error of each outer weight. A resampled dataset was created from the original dataset of

N participants by randomly selecting N participant rows from this set whilst allowing a row to be selected more than once (bootstrapping with replacement). The PLS model was fitted and orthogonal Procrustes rotation was applied to correct for sign-flipping by rotating each set of bootstrapped outer weights for each dataset and component onto the equivalent set for the first imputed dataset. This was repeated 500 times and for each imputed dataset.

The bootstrap means that the subjects contributing to each PLS model were varied, allowing us to estimate the reliability of each weight based on how it varied from sample to sample. The standard errors were calculated for each variable weight across these 500 samples for each imputed dataset. Then the average outer weights across imputations were calculated and standard errors were pooled according to Rubin's rules (which also accounts for the additional variance due to imputation). The averaged weights and their corresponding pooled bootstrap standard errors were used to calculate the t-value (ratio of weight to standard error) in order to assess the contribution of each variable to the PLS latent variable. The variables in each dataset with a t-value above the critical t-value (e.g. 1.985 for a two-tailed t-test at a significance level of 0.05 with 94 participants) were considered to load reliably onto the latent variables.

2.3.4 Mediation Analysis with Multi-block PLS

RGCCA can be used to find the set of latent variables for each component that best describe the relationship between multiple datasets simultaneously. For example, in the case of three datasets (X_X, X_M and $X_{Y,}$), latent variables are found that maximise:

$$cov(L_X, L_M) + cov(L_M, L_Y) + cov(L_X, L_Y)$$

These latent variables can then be submitted to mediation analysis, to assess whether dataset L_M mediates the relationship between L_X and L_Y, as shown in the path model in Figure 7. In particular, we wish to examine whether aspects of either the child's environment or measures of their development partially mediate the relationships found between the datasets. This provides a way to assess the relative strength of the relationships between the datasets and to tease apart the potential mechanisms by which the environmental, cognitive and neural factors may exert an effect on the academic and behavioural outcomes. In this context, this is calculated by linear regression of the set of latent variables (for a given component), as in Path Modelling, as follows:

$$L_M = aL_X$$

$$L_Y = c'L_X + bL_M$$

The strength of the mediating effect is given by $ab = a \times b$ (the indirect path coefficient, which is equal to the difference in direct path coefficient before and after the mediating variable is added: $c - c'$). Therefore if ab is significant, L_M is found to have a significant mediating effect on the relationship between L_X and L_Y (MacKinnon et al., 1995). Note that in this study, the latent variables were standardised before calculating the path coefficients, and therefore all of the path coefficients reported are standardised.

1)
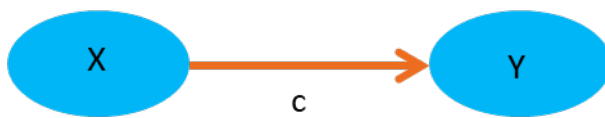
2)
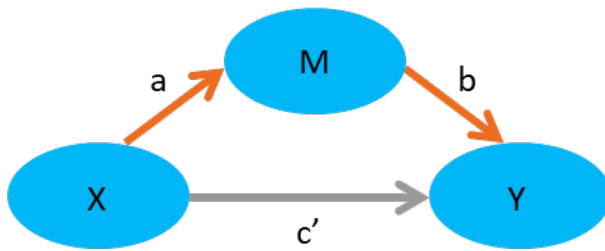
Figure 7 Example of the path models used for the mediation analysis

2.3.5 Assessing significance and reliability of mediation effect

Whilst there are methods available for assessing the significance of a mediation effect via permutation testing similar to the two-block case, it has been shown that these have a tendency towards type I errors and significance based on bootstrapping of the indirect path coefficient, ab, is recommended instead (Taylor and MacKinnon, 2012). In addition, the indirect effect does not typically follow a normal distribution as it is the product of two path coefficients. Therefore, it is also recommended to calculate the asymmetric confidence intervals from the distribution of indirect effects from each bootstrap, such as the percentile confidence intervals (Nitzl et al., 2016), in contrast to the symmetric confidence intervals found using the standard error and t-distribution. Note, that it is also recommended that non-bias corrected confidence intervals are used as it has been shown that bias correction can also inflate type I error rates (Taylor and MacKinnon, 2012). However, research into methods for finding bootstrap confidence

intervals in the case of both multiple imputation and non- normal data is still inconclusive (Rubin's rules for confidence intervals assume normality). In 2016, Nitzl et al. validated a method in which the bootstrapped indirect effects for each imputed dataset are pooled to create one large distribution and the asymmetric percentile confidence intervals were calculated using this. Therefore, in this study we have used both this method to calculate the confidence intervals in addition to the more commonly used method in PLS path analysis that uses the bootstrap standard error to calculate the t-value and compares it to the critical value.

The significance of the indirect effect and the reliability of the outer weights for each latent variable and the path coefficients are assessed using 500 bootstrap samples drawn with replacement. The PLS model was fitted for each sample and for each imputation and orthogonal Procrustes rotation was applied to correct for sign-flipping by rotating each set of bootstrapped outer weights and latent variables for each dataset and component onto the equivalent set for the first imputed dataset. Latent variables were scaled to unit variance and submitted to the path analysis equations to calculate the path coefficients and indirect effect for each bootstrap sample.

95% confidence intervals were found for the distribution of the indirect effects from the combined bootstrap samples for each imputation. If the upper and lower confidence intervals did not include zero between them, then the mediation effect is found to be significant. To assess reliability of the outer weights and the path coefficients, results were combined across imputations using Rubin's rules and the averaged outer weights and the path coefficients and their corresponding pooled bootstrap standard errors were used to calculate the t-values. The variables in each dataset with t-value above the critical t-value (e.g. 1.985 for a two-tailed t-test at a significance level of 0.05 with 94 participants) were considered to be reliable estimates.

2.4 Feature Selection

Whist the PLS methods are useful in characterizing the relative importance of each variable to the outcomes, it is also interesting to consider which features are the most important to measure in order to best predict the outcomes. Note that the high correlations between some of the variables imply that measurement of some variables may be redundant as the information is captured elsewhere. In order to do this, we applied the least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1994) to perform both feature selection and regularization in the regression of the datasets

onto each outcome. LASSO regularization, although a relatively recent method, has rapidly emerged as one of the most prominent and widely used methods of feature selection, demonstrated to outperform many classic methods such as stepwise and backwards selection. It is a type of linear regression that applies L1 regularisation to shrink some of the regression coefficients towards zero, reducing the original set of predictor variables in order to create a simple and sparse structure.

This is achieved as follows:

Let Y be the dependent variable to be predicted from a set of independent variables X_j

$$Y = \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_n X_n + \epsilon$$

In ordinary least squares regression, the regression coefficients, β_j, are predicted by minimising the 'sum of squares' error term given by:

$$min \sum_{i=1}^{n} \left(Y_i - \sum_j \beta_j X_{ij} \right)^2$$

Where Y_i and X_{ij} are the variable values for the ith observation.

L1 regularization is performed by adding the summed absolute magnitude of all β_j coefficients, weighted by a tuning parameter, λ, to the sum of squares error and minimizing this instead:

$$min \sum_{i=1}^{n} \left(Y_i - \sum_j \beta_j X_{ij} \right)^2 + \lambda \sum_j |\beta_j|$$

Note that as $\lambda \to 0$, this tends towards ordinary least squares regression, and as λ is increased, the regularization is stronger and fewer variables are selected. LASSO Regression is repeated for several different values of λ, and a model can be selected based on a number of criteria. In this study, the model was selected based on the lowest Bayesian information criterion (BIC) value, a criterion that aims to identify a model's quality based on a trade-off between goodness of fit and model complexity. An additional analysis using the Akaike Information Criterion is included in Appendix D for comparison.

2.4.1 MI-LASSO

If LASSO regression is applied to each MI dataset separately, it is possible that different variables may be selected for each imputation. In a recent paper, Chen and Wang proposed a novel method for analysing the MI datasets jointly to ensure that the variables selected remained the same across datasets and demonstrated it to have better sensitivity, specificity and mean squared error in comparison to alternative methods such as stepwise variable selection for multiply imputed data (Wood et al., 2008) and LASSO on complete cases (Chen and Wang, 2013). In this, they applied the group LASSO (Huang and Zhang, 2009; Yuan and Lin, 2006), designed to be used when covariates can be separated into groups of similar variables, by treating the same variables across the imputed datasets as a 'group'.

If we have m imputed datasets, let Y_k, $\beta_{k,j}$ and $X_{k,j}$ be the set of variables and coefficients for the kth imputation. In order to apply the group LASSO, the term to minimize is adjusted to ensure that the $\beta_{1,j}$, $\beta_{2,j}$, $\beta_{3,j}$ ... $\beta_{m,j}$, for the jth variable don't vary greatly between the m sets and ultimately that all are either zero or non-zero, resulting in the same variables being selected across imputations for each value of λ:

$$min \sum_{k=1}^{m} \sum_{i=1}^{n} \left(Y_{k,i} - \sum_{j} \beta_{k,j} X_{k,ij} \right)^2 + \lambda \sum_{j} \sqrt{\sum_{k=1}^{m} \beta_{k,j}^2}$$

In summary, this study aims to shed light on the bigger picture relationship between the environment and a child's academic attainment, behaviour, cognition and functional and structural brain development. It expands on previous studies by investigating how multiple aspects of a child's environment combine to create an environment conducive to positive child development, using partial least squares methods. Furthermore, multi-block PLS is used to explore the possible pathways between the environment and academic and behaviour outcomes by considering the factors that mediate these relationships. Finally, LASSO feature selection is used to identify the environmental factors that are most important to measure for the prediction of child academic and behaviour outcomes. Each aspect of child development will be considered in turn in the following chapters, beginning with the academic and behaviour outcomes, followed by cognition, structural connectome and finally the functional connectome. Each adds a piece of evidence relating to the child environment-development relationship that will be brought together and discussed in the final section.

3 ENVIRONMENT AND CHILDHOOD OUTCOMES

3.1 Introduction

The academic ability and behaviour of children are key markers of child development and are typically the target outcomes of interest for practitioners and policy makers. These childhood outcomes are considered integral to a host of lifelong positive outcomes such as wellbeing, life satisfaction, employment, income and success (e.g. see (Feinstein, 2000)). As such, identifying the factors in a child's environment that are most strongly associated with these outcomes remains a key research task, important for supporting positive development in children.

Key aspects of the environment, particularly standard markers of SES such as income, occupation and education, have been consistently associated with these outcomes. However, the environment cannot be boiled down to simplistic single measures and yet little has been done to capture the complexity of a child's environment. In particular, we wish to address this by investigating how the complex and often related set of environmental factors that make up a child's environment combine to produce an environmental profile that is predictive of positive academic and behavioural outcomes. PLS methods are particularly well designed to model relationships between datasets containing multiple, potentially collinear measures. We will use these methods to identify the sets of environmental factors that most strongly covary with the academic and behavioural outcomes.

In addition, reducing the outcome attainment gap due to the relationship between children from different SES backgrounds remains a key goal of practitioners and policy-makers. It is likely that this attainment gap is mediated by factors in the child's environment. SES may have an impact on these intermediate factors which in turn impact child development outcomes. It is these mediators that are often targeted by intervention studies. Consequently, providing a better understanding of these relationships will help to develop more effective and efficient interventions. We

therefore aim to use three-block PLS to investigate the relationships between standard SES measures (income, occupation and education), other environmental factors and academic ability and behavioural outcomes. In short, we want to understand how the wider environment mediates the relationship between SES and academic ability and behaviour.

Furthermore, it is generally impractical, particularly for practitioners, to collect a large amount of data on each child, such as the questionnaires used in this study. It would, therefore, be useful to identify a smaller set of variables that can be combined to provide a good prediction of the outcomes. LASSO feature selection provides a method for achieving this by selecting the set of predictor variables that explain the most variance in the dependent variable whilst forcing the regression coefficients of less important variables to zero. It is also particularly useful when predictor variables are highly correlated, as is likely to be the case in the questionnaires used, resulting in the measurement of redundant information. LASSO selects the variables amongst a highly correlated set that best capture the shared variance, whilst removing redundant variables from the regression whose variance is better explained by the selected variables.

3.2 Methods

3.2.1 Participants and tasks

This analysis used the sample of 94 children who had less than a 15% of their data missing (45 males, mean age 9:10 and age range 6:11-12:9). Academic ability was measured using three reading subtests and two maths subtests from the Woodcock Johnson III Form B Tests of Achievement (McGrew et al., 2007; Woodcock et al., 2001). The Letter-Word Identification, Passage Comprehension and Reading Fluency subtests were used to calculate an age-standardised reading score and the Calculation and Maths Fluency subtests were used to calculate an age-standardised maths score using the Woodcock Johnson III Normative Update Test. Behaviour was measured using the BRIEF (Baron, 2000) and SDQ (Goodman, 1997) questionnaires completed by the primary caregiver. The 8 subscales from the BRIEF (Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of and Monitor) were combined in a factor analysis and the scores for the first factor were used as a summary measure of behavioural difficulties related to executive functions deficits. The 5 subscales from the SDQ (hyperactivity, emotional symptoms, conduct problems, peer

problems and prosocial behaviour) were submitted to a factor analysis and the first component was used to provide a summary score relating to externalising and internalising behaviour. Note that only the first factor was selected for both the BRIEF and SDQ questionnaires based on the eigenvalue > 1 criteria, inspection of the scree plot and parallel analysis. BRIEF and SDQ scores were multiplied by -1 so that a higher score relates to better behaviour. Finally, the 21 factors from the factor analyses of the questionnaire data were used to characterize the child's environment. Please see Chapter 2 for more detail and Appendix B for the domains and corresponding questionnaire items.

3.2.2 PLS between the environment and child outcomes

As an initial way to visualize the relationships, the Pearson correlation between each of the 2 academic, 2 behaviour and 21 environmental factors was first calculated. However, considering 600 pairwise correlations is not only difficult to interpret but would present a problematic challenge in correcting for multiple comparisons. In addition, the environmental variables are related to each other and so their relationship to the outcome measures will not be independent for each variable but rather consist of a profile of environmental variables that move together in the direction of the outcomes. Thus, PLS was used to draw out the overall profiles of environmental factors that most strongly covary with either the academic or behavioural measures.

The 'PLS' function from the MixOmics Package (Rohart et al., 2017) and a series of in-house scripts developed for this project in R (available at https://github.com/ajohnson62) were used to apply 2 Block Canonical PLS (RGCCA with new Mode A and Horst's scheme function) to investigate the relationships between the environmental domains and the outcomes. Each variable was standardized by subtracting the mean and dividing by the standard deviation. The environmental variables were used as dataset X and either the two academic variables or the two behavioural variables were used as dataset Y resulting in two PLS models. PLS was applied to each imputed dataset in turn and the key results were then pooled across imputation using Rubin's rules as described below.

The significance of each pair of latent variables extracted (2 in each case due to there being 2 outcomes each time) was tested by repeatedly fitting a PLS model under 1000 permutations of the questionnaire data with respect to the outcome data. The permuted samples were kept consistent across imputations. P values for each imputation were

calculated by comparing the covariance between each pair of latent variables, $Cov(L_{Xi}, L_{Yi})$, in the original dataset to the null distribution from the permuted data. These were pooled across imputation to obtain an overall p value for the significance of each component, according to the Licht-Rubin method or averaging z-transformed p values before transforming back into 0-1 space (Licht, 2010).

The stability of each variable weight onto a given latent variable was assessed for each imputed dataset using 500 bootstrap samples selected with replacement. A new PLS model was fitted to each sample to obtain a distribution of weights for each variable. Each set of outer weights was aligned with the original sample for the first imputed dataset using Procrustes rotation to account for the sign flipping. The resulting weight distributions were used to calculate the mean and standard errors for each weight for each imputed datasets. The results were then pooled across imputation using Rubin's rules. The average mean weight across imputations was calculated for each variable. The total standard error for each imputation was combined using Rubin's standard error formulas to account for the additional variation due to imputation of missing data. The average outer weights and standard errors were then used to calculate the t-value (ratio of weight to standard error). The variables in each dataset with t-value above the critical t-value of 1.985 (two-tailed t-test at a significance level of 0.05 with 94 participants) were considered to load reliably onto the latent variables. Each outer weight is proportional to the covariance between the individual variable in one dataset and the latent variable in the other dataset and so indicates the importance of each variable to describing the relationship.

3.2.3 PLS mediation

In order to investigate whether the wider environment mediates the relationship between standard SES measures and the outcomes, we applied 3-block Canonical PLS (RGCCA with new Mode A and Horst's scheme function) using the 'block.pls' function from the MixOmics Package (Rohart et al., 2017) and a series of in-house scripts developed for this project in R (available at https://github.com/ajohnson62). As in the two-block case, each variable was standardized by subtracting the mean and dividing by the standard deviation. Parent income, education and occupation were used as the standard measures of SES in dataset X. As before, either the two academic variables or two behaviour variables were used as dataset Y. Primary caregiver's hours work per week was excluded as it might be considered a proxy measure for equivalised income and the

remaining 18 environmental domains were used as the third dataset, M. Block PLS results in sets of 3 latent variables, one for each of the datasets that best explain the covariance between the three datasets. The latent variables for each imputed dataset were rotated by 1 or -1 to align them with the variables from the first imputed dataset and standardised to unit variance.

In order to assess the mediating effect of the other environmental domains, the sets of latent variables for each imputation were submitted to a path analysis as shown in Figure 8. The indirect path coefficients 'ab' were calculated and the average across imputation calculated.
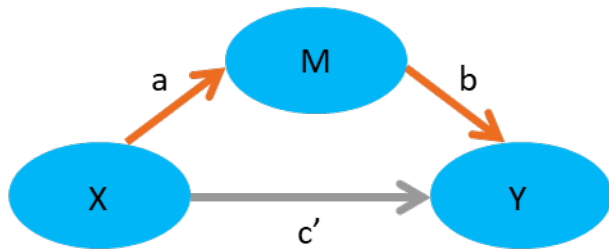
Figure 8 PLS mediation analysis: path model example

The significance of the indirect effect and the reliability of the outer weights for each latent variable and the path coefficients was tested using 500 bootstrap samples with replacement. The sign of each latent variable was flipped according to Procrustes rotation to align it with the original latent variable for the first imputed dataset and standardised before calculating the path coefficients. 95% confidence intervals were found for the distribution of the indirect effects from the combined bootstrap samples for each imputation. If the upper and lower confidence intervals did not include zero between them, the other environmental variables were considered to (partially) mediate the relationship between SES and the outcomes. The reliability of the outer weights and the path coefficients were found by combining results across imputations using Rubin's rules and calculating the t-values. The variables or path coefficients in each dataset with t-value above the critical t-value were considered to be reliable estimates.

3.2.4 MI-LASSO feature selection

The MI-LASSO method for multiply imputed data was applied to identify the most important predictors amongst the set of environmental variables for predicting each of the outcomes in turn. The 21 environmental domains were used to predict (separately) reading, maths, BRIEF and SDQ scores resulting in 4 regression models. The regression

was iterated over 102 values of the tuning parameter, λ, from 0 to 128 and the optimal model was selected by using the Bayesian information criterion (BIC). BIC is a relatively stringent criterion which selects only a few variables for inclusion in the model and so a supplementary analysis was conducted using an alternative criterion, the Akaike information criterion (AIC) in order to select a slightly larger set of variables. All variables were demeaned and scaled to unit standard deviation before fitting the model to ensure that the regression coefficients obtained are standardized.

3.3 Results

3.3.1 Pearson correlation results

The Pearson correlation between each of the factor scores and the other measures was calculated and can be seen in Figure 9. There is a relatively strong correlation between the two academic outcomes ($r=0.56$) and between the two behavioural outcomes ($r=0.68$). In general, there are positive correlations between all of the variables with the notable exception of discipline and technology use which correlate negatively with almost all other measures including the outcome measures. Note that many of the correlations between the environmental measures demonstrate collinearity across domains which suggests that methods that are designed to work under this condition, such as PLS methods and LASSO, are necessary. In order to focus on the relationship between the environment and the outcomes, Figure 10 shows an enlarged plot of these. Note that the questionnaire items that make up each environmental domain can be found in Appendix B.

Figure 9 The Pearson correlation between the environment domains and the child outcomes

	Reading.ability	Maths.ability	BRIEF	SDQ
Equivalised.income	0.01	0	0.12	0.16
Caregiver.education.av	0.25	0.18	0.21	0.23
Caregiver.occupation.av	0.2	0.14	0.14	0.14
Subjective.SES	0.03	-0.14	0.02	0.08
Number.caregivers	0.07	0.11	0	0.01
Number.siblings	0.07	0.27	0.07	0.06
Primary.caregiver.hours.work	0.12	0.01	0.15	0
Primary.caregiver.skills	0.1	-0.12	0.27	0.18
Acquaintance.skills	0.14	0	0.15	0.21
Neighbourhood.SES	0.11	-0.04	0.19	0.16
Attitude.child.education	0.26	0.24	0.24	0.19
Child.health	0.2	0.11	0.3	0.4
Primary.caregiver.wellbeing	0.03	-0.06	0.2	0.18
Time.with.family.and.friends	0.02	-0.09	0.1	0.08
Opinion.family.relationships	-0.03	-0.13	0.21	0.27
Childcare	0.2	0.1	-0.07	0.1
Technology.use	-0.25	-0.23	-0.18	-0.29
Discipline	-0.19	-0.02	-0.51	-0.47
Other.language.use	0.01	-0.02	0.1	0.05
Reading.at.home	0.33	0.25	0.28	0.33
Attitude.to.neighbourhood	-0.08	0.02	0.04	0.16
Rules.and.chores	-0.01	-0.08	-0.05	-0.12

Figure 10 The Pearson correlation between the environment domains and child outcomes: enlarged plot

3.3.2 PLS results

3.3.2.1 PLS: What is the environmental profile that most strongly relates to academic outcomes?

The first pair of latent variables was found to be significant for explaining the relationship between the environmental domains and the academic outcomes (p = 0.012, explained variance: 15% environment and 77% academic scores, average correlation between environment and academic latent variable = 0.46). The second component was not significant. Maths and reading both loaded strongly onto the first component (see Figure 11). Reading at home and attitude to education, both measures that are indicative of attitudes to child academic development, show a strong positive relationship to the academic outcomes. Technology use, on the other hand, has a relatively strong negative

relationship. Parent occupation and education, number of siblings, child health and childcare show a moderate positive relationship to academic outcomes. It is particularly interesting to note that income has no relationship to academic outcomes in this sample.
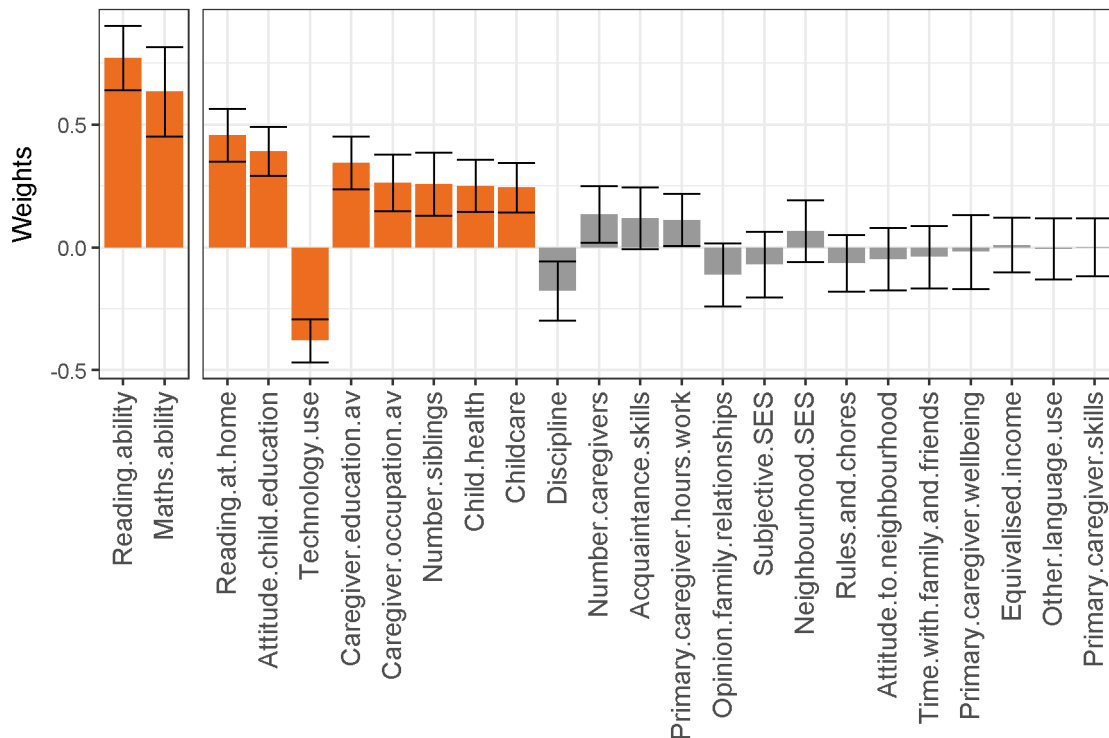
Figure 11 The environment and academic ability: outer weights for the first latent variables, averaged across imputations. Error bars are the standard errors obtained from the bootstrap samples, pooled across imputations using Rubin's rules. Reliable loadings (according to the critical t-value) are coloured orange.

Because maths and reading load relatively equally onto the academic latent variable, this provides us with a summary measure of academic ability that is maximally related to the environmental profile. We can use this to group children into high and low academic ability groups based on a median split of the first academic latent variable to illustrate how the environmental profiles of children vary depending on individual academic ability. Figure 12 compares the average standardised environmental scores between children of low and high academic ability. The ability groups were based on the median split of the academic latent variable so that children with scores above median were in the high academic ability group and vice versa. This figure helps to visualise how environmental profiles for children differ for those that have lower or higher academic ability. It also demonstrates the first PLS component is selecting

environmental variables that are maximally different between low and high ability groups.
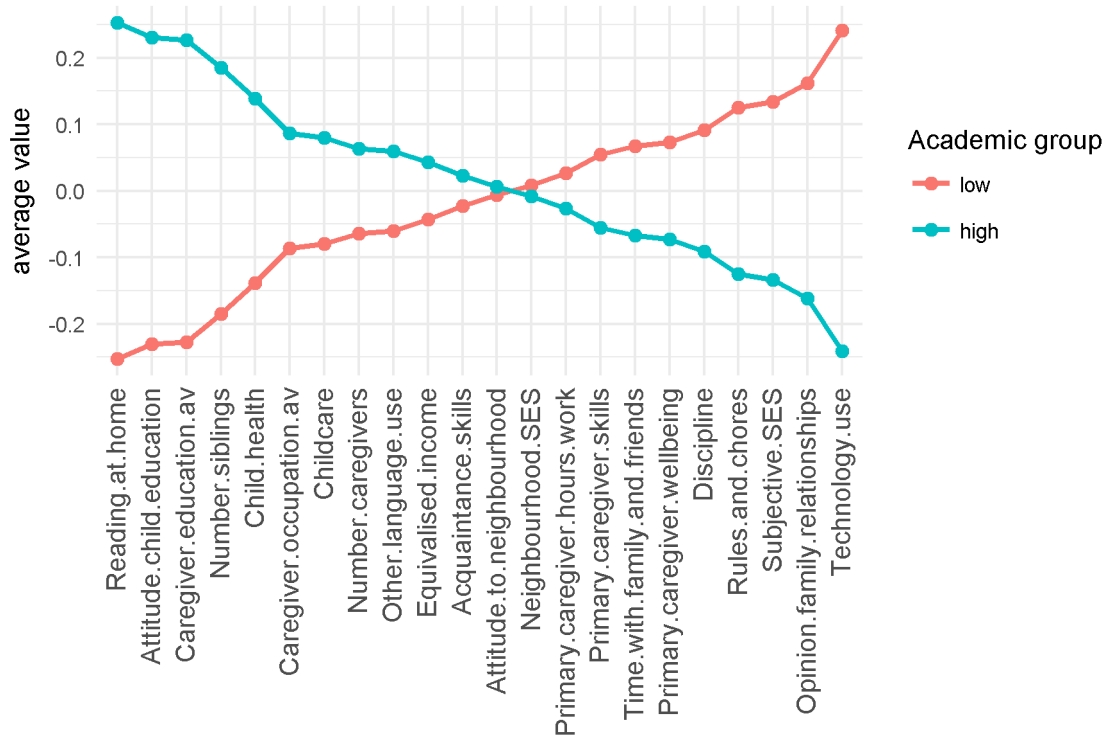
Figure 12 The environment and academic ability: average score for each environmental domain, grouped by academic ability

3.3.2.2 PLS: What is the environmental profile that most strongly relates to behaviour outcomes?

The first pair of latent variables was found to be significant for explaining the relationship between the environmental scores and the behavioural outcomes (p< 0.001, explained variance: 16% environment and 84% behavioural, average correlation between environment and behaviour latent variable = 0.58). BRIEF and SDQ both loaded strongly onto this component (Figure 13). Discipline, a measure of the type and frequency of discipline the child receives is strongly negatively related to behaviour indicating that the greater the discipline, the poorer the behavioural outcomes. Technology use is also inversely related to child behaviour. Child health, reading at home, and, to a lesser extent, subjective family relationship quality, social resources (both primary caregiver and acquaintance skills), attitude to child education, parent wellbeing and neighbourhood SES all have a positive relationship with the behavioural outcomes. Much like the results for the academic outcomes, parent education shows a moderate positive relationship with behaviour but there is no relationship between income, subjective SES or parent occupation and the behaviour outcomes.
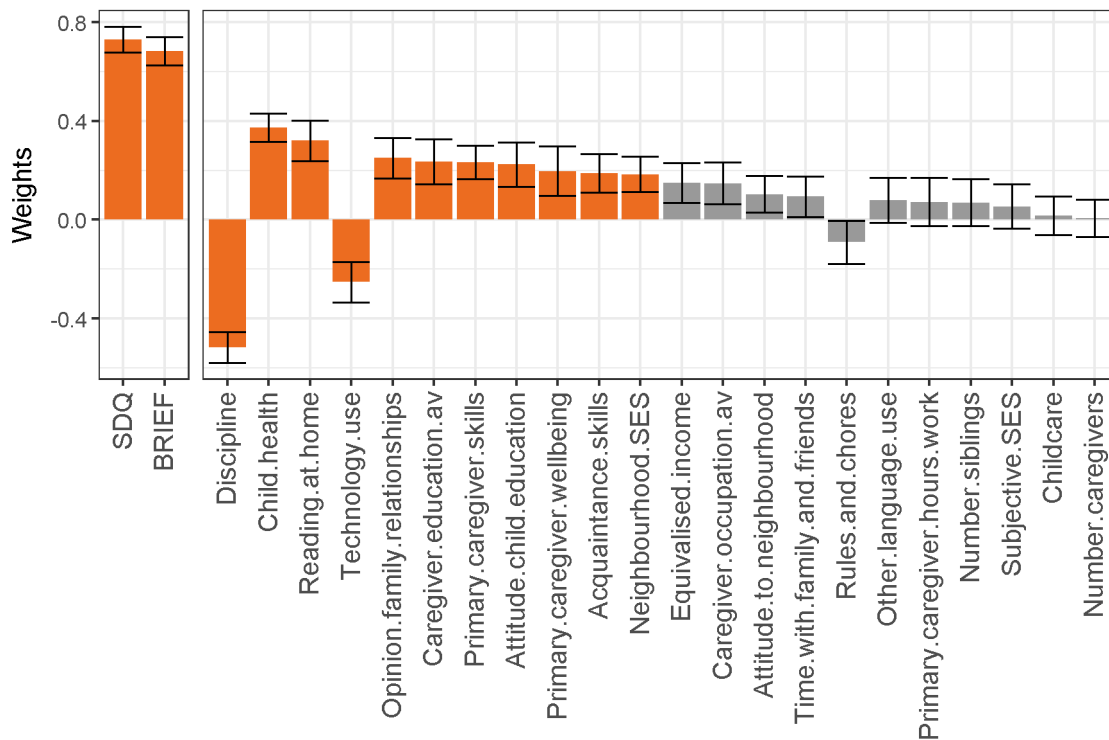
Figure 13 The environment and child behaviour: outer weights for the first latent variables, averaged across imputations.

As for the academic model, we use the behaviour latent variable to illustrate the difference in environmental profile between children that have higher or lower behavioural scores according to the median split in Figure 14. The environmental scores were standardised to allow comparison between environmental domains. This figure demonstrates that the environmental profiles for children differ for those that have lower or higher behaviour outcomes. In general, it also demonstrates the first PLS component is selecting environmental variables that are maximally different between low and high behaviour groups. However, it must be noted that this median split of the data is for illustrative purposes only, and it is the outer weights in Figure 13 that are proportional to the covariance (and correlation) between each environmental variable and the behaviour outcome and therefore illustrates the importance of each environmental variable in relating to behaviour.
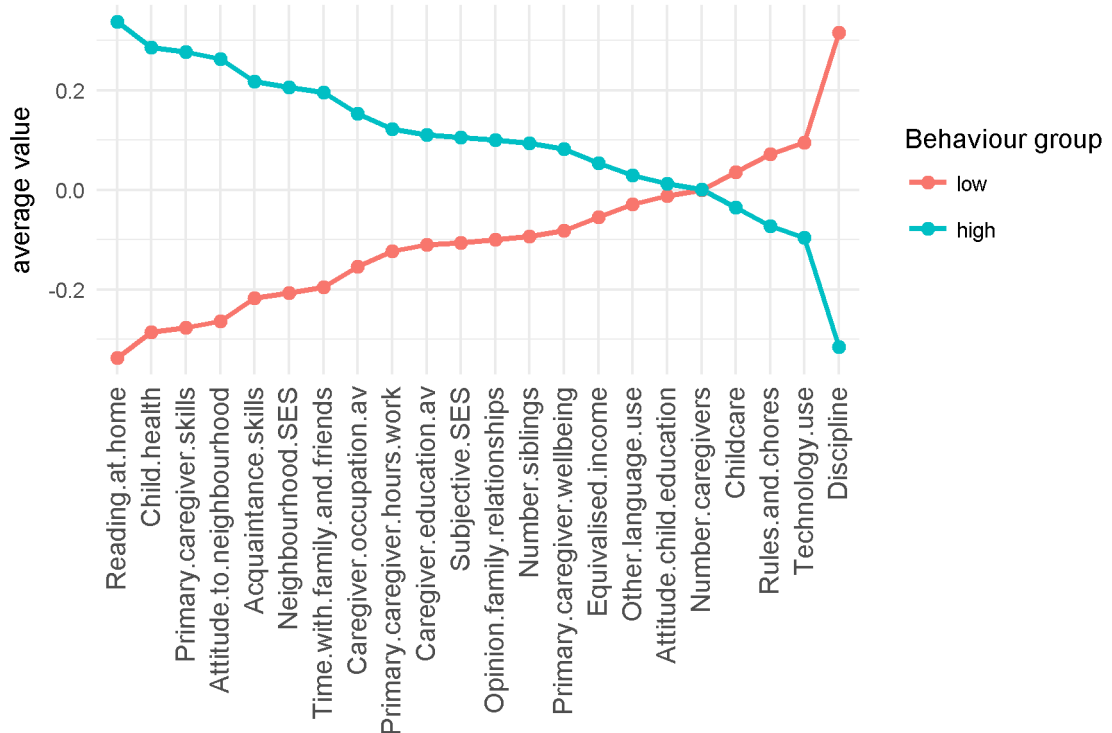
Figure 14 The environment and child behaviour: average score for each environmental domain, grouped by behaviour

3.3.3 PLS mediation results

3.3.3.1 Does the environment mediate the relationship between SES and academic ability?

The mediating effect of the wider environment on the SES-academic relationship was significant for the set of latent variables in the first component (indirect path coefficient ab=0.28±0.10, 95% CI [0.16,0.56]). The path model can be seen in Figure 15 and the variable weights can be seen in Figure 16. The second component was not found to be significant.
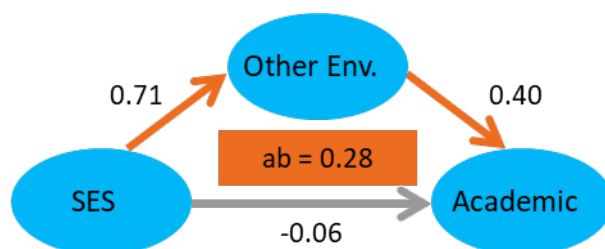
Figure 15 PLS mediation analysis between SES, the wider environment and academic ability: the path model

Both reading and maths load onto the component, though the maths outer weight is lower than reading. Each of the three SES measures load onto the component, with income having a lower outer weight than education and occupation. Finally, a number of environmental domains reliably relate to this mediating effect. Specifically, in order of magnitude: inverse technology, childcare, attitude to child education, reading at home, child health, use of other languages, number of caregivers, inverse discipline, primary caregiver personal skills in the resource generator subscale and inverse time with family and friends. In the case of three datasets, the magnitude of each of these variable weights relate to the sum covariance between each variable in one dataset and the latent variable of each of the other datasets. This means that, as in the two-block case, the higher the weight, the more important this variable is for explaining the three-way relationship.
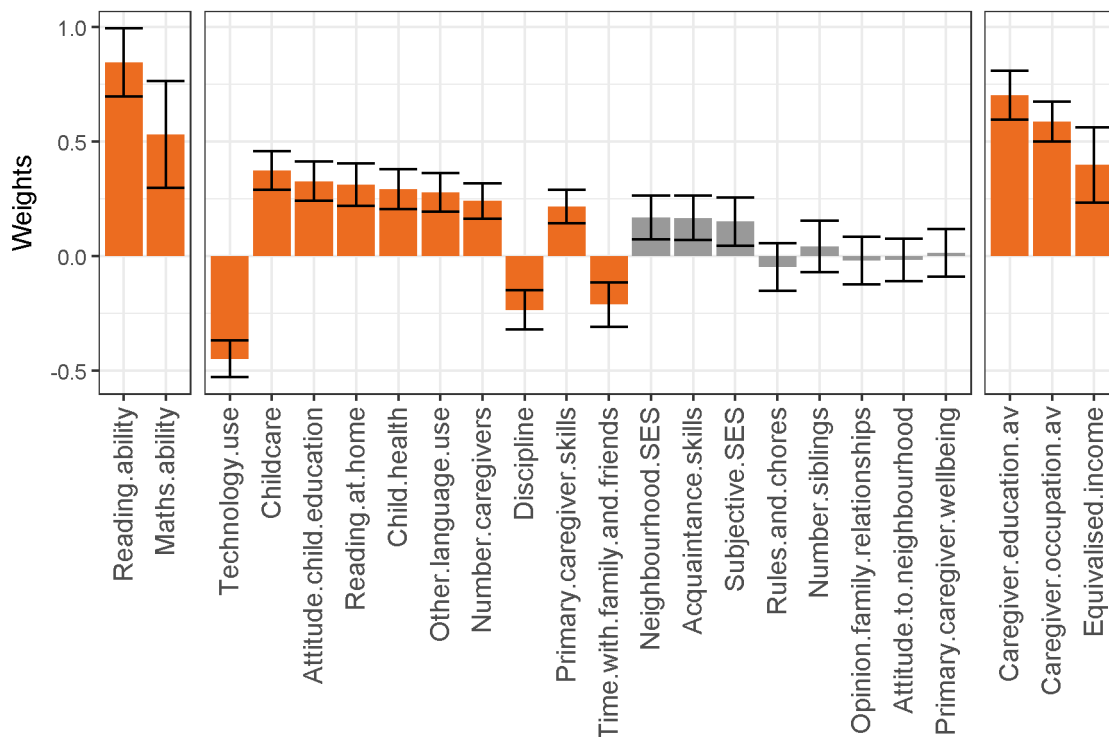
Figure 16 PLS mediation analysis between SES, the wider environment and academic ability: outer weights for the first latent variables, averaged across imputations

This relationship is illustrated in Figure 17, in which the SES latent variable is plotted against the Academic latent variable, showing that higher SES relates to higher academic outcomes. Furthermore, the participant scores are coloured according to their environment score based on a median split of this score. It can be seen that higher academic and higher SES both relate to higher environmental scores and vice versa,

which relates to the mediation effect found. This visualisation highlights that children from higher SES backgrounds have an environmental profile that differs from those of a low SES background, particularly in the domains listed above, that in turn relates to their higher academic outcomes.
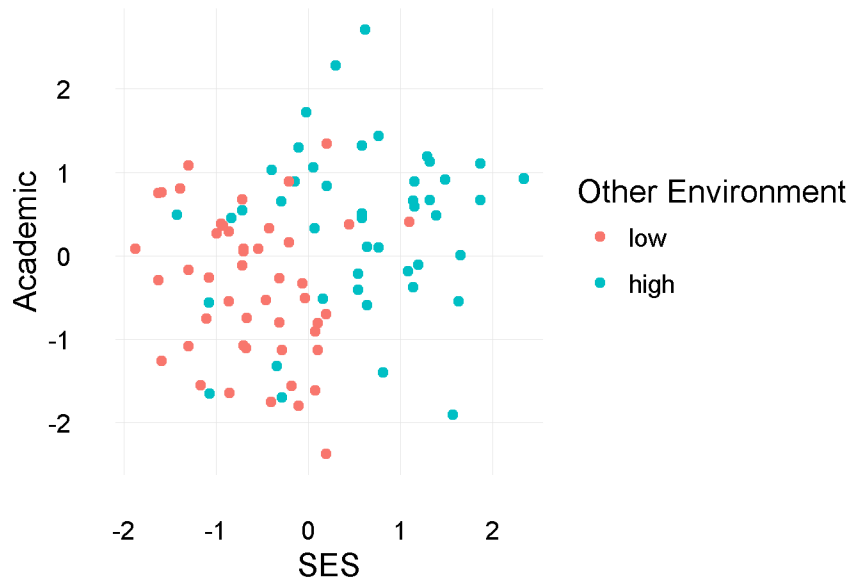
Figure 17 PLS mediation analysis between SES, the wider environment and academic ability: scatter plot illustrating the relationship between the SES and academic latent variables, coloured by median split of the wider environment domain

3.3.3.2 Does the environment mediate the relationship between SES and behaviour outcomes?

The mediating effect of the wider environment on the SES-behaviour relationship was significant for the set of latent variables in the first component (indirect path coefficient ab=0.42±0.08 95% CI [0.32,0.64]). Note that the original correlation between SES and behaviour, before accounting for the wider environment, was reliable but not very strong (r= 0.24±0.1 95% CI [0.04,0.43]). The path model can be seen in Figure 18.
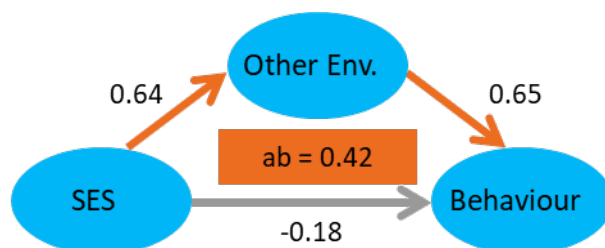
Figure 18 PLS mediation analysis between SES, the wider environment and behaviour: path model

Both BRIEF and SDQ load onto the component. Each of the three SES measures load evenly onto the component. Finally, a number of environmental domains reliably relate to this mediating effect. Specifically, in order of magnitude: inverse discipline and technology, child health, primary caregiver personal skills, use of other languages, attitude to child education, reading at home, childcare, neighbourhood SES, subjective SES, acquaintance skills in the resource generator and number of caregivers. The second component was not significant.
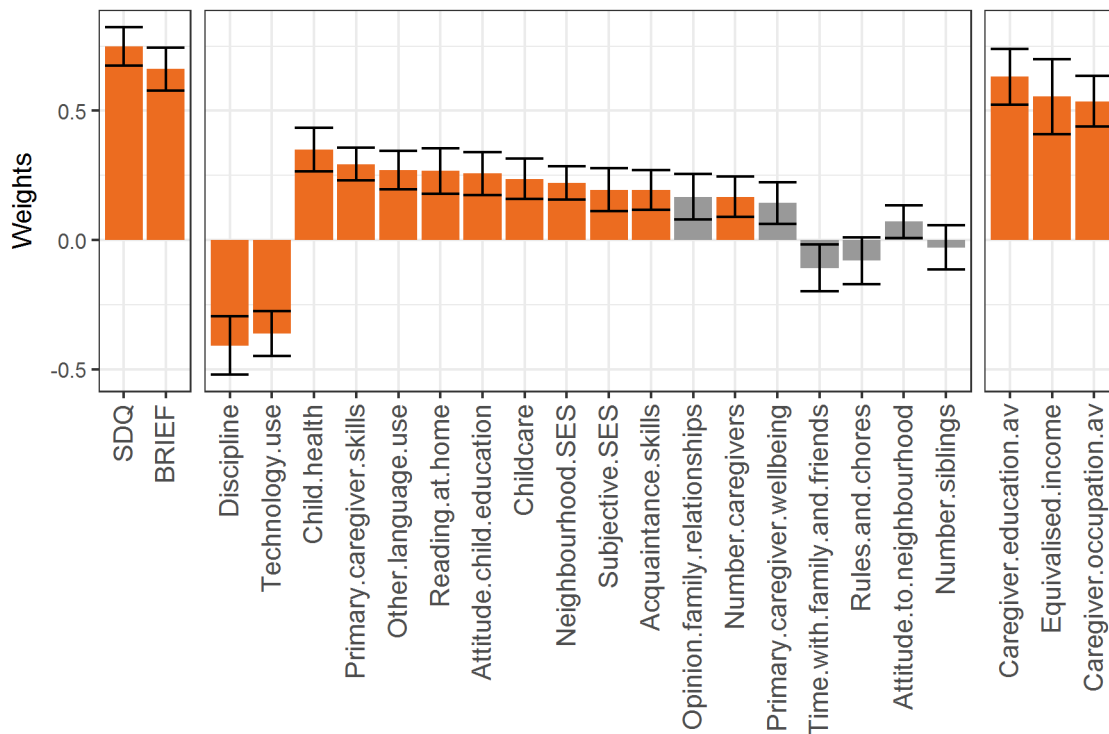
Figure 19 PLS mediation analysis between SES, the wider environment and behaviour: outer weights for the first latent variables, averaged across imputations

The SES latent variable is plotted against the behaviour latent variable in Figure 20, illustrating the positive relationship between SES and behaviour. As in 3.3.3.1, the participant scores are coloured according to their environment score based on a median split of this score. Higher behaviour scores and higher SES both relate to higher environmental scores and vice versa. This illustrates that children from higher SES backgrounds have an environmental profile that differs from those of a low SES background that in turn relates to their higher behaviour outcomes.
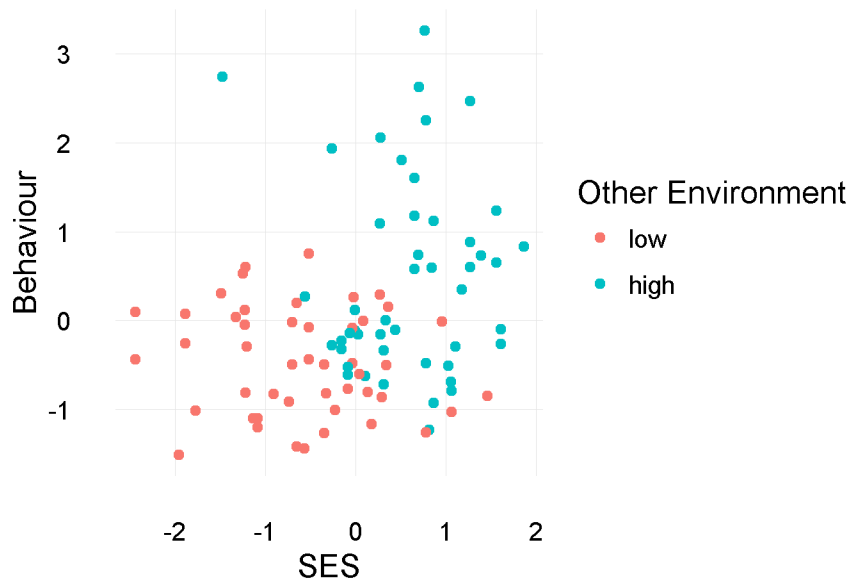
Figure 20 PLS mediation analysis between SES, the wider environment and behaviour: scatter plot illustrating the relationship between the SES and behaviour latent variables, coloured by median split of the wider environment domain

3.3.4 MI-LASSO feature selection results

Figure 21 and Figure 22 demonstrate the LASSO regression for predicting maths scores. The left side of Figure 21, where $\lambda = 0$ gives the regression coefficients that would be found using ordinary least squares (OLS) regression. As λ is increased, a greater penalty is applied to the sum of regression coefficients, causing them to be reduced, and ultimately resulting in some coefficients being forced to zero. The variables that are the least predictive of maths drop out quite quickly, whilst the variables that explain the greatest amount of unique variance in the outcome, such as number of siblings, reading at home and attitude to child education remain in the regression as λ increases. In Figure 22, the models are ranked in order of decreasing BIC and the final model selected can be seen on the right-hand side of the plot.
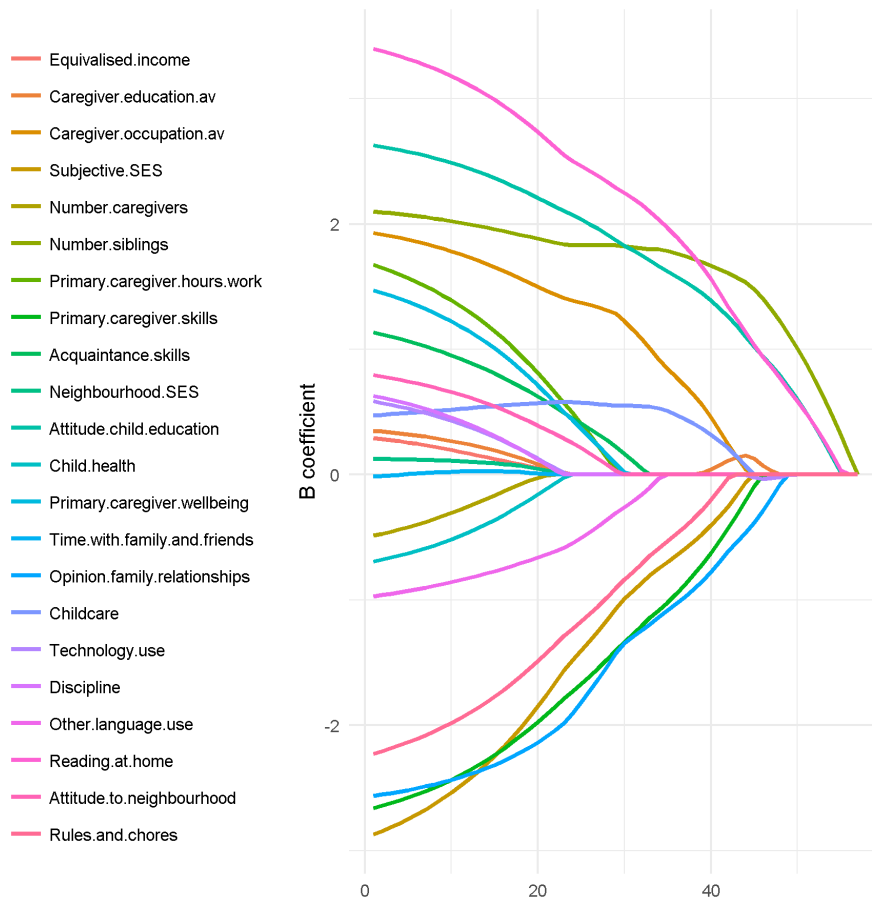
Figure 21 LASSO prediction of maths scores: the regression coefficients for each variable are plotted on the y axis as the tuning parameter, λ, is increased.
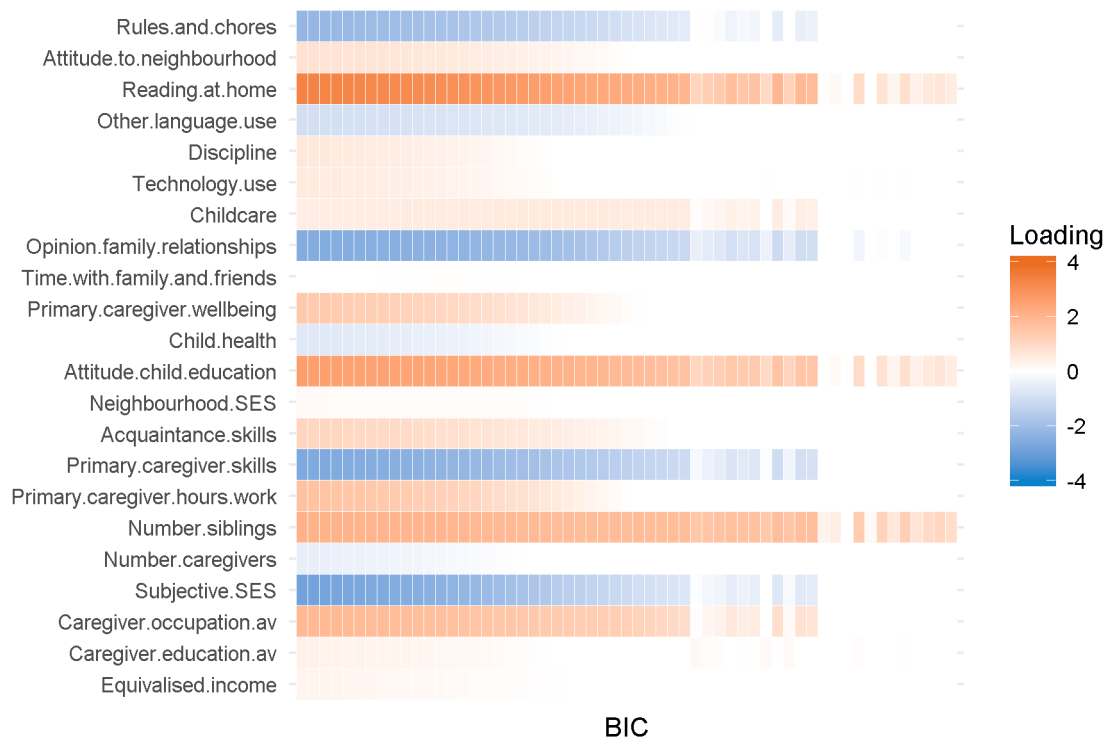
Figure 22 LASSO prediction of maths scores: regression coefficients for each of the LASSO modes, ordered by decreasing BIC . The model selected as optimal is given on the right of the plot.

The average regression coefficients that would be found using OLS regression for each outcome, averaged over imputations, can be seen in Figure 23. The large number of relatively high regression coefficients demonstrates the need for LASSO selection. Note that most of the p values for each coefficient are not significant after correcting for multiple comparisons due to the large number of variables (p <0.002 is needed to reach significance under Bonferroni correction). In addition, several of the regression coefficients can be seen to behave strangely, having signs that are in the opposite direction to the correlation between that predictor variable and the outcome, due to multicollinearity between the predictor variables.

In contrast, Figure 24 shows the regression coefficients for the optimal LASSO regression model, selected based on the BIC values. Under the BIC criterion, 1 – 4 variables are selected as the best set of predictors for each outcome. As found using PLS, reading at home is selected as an important predictor for each of the academic and behaviour outcomes. In addition, attitude to education and the number of siblings are also selected for the prediction of maths. Child health is selected for predicting both behaviour scores, although the loading is very low for BRIEF. Finally, acquaintance

skills (Resource generator scale) are selected for predicting SDQ scores. None of the standard SES measures are selected as important predictors for any of the outcomes, indicating that the covariance between these and the outcomes is well described by variance in the other selected environmental variables. This corroborates the findings of the PLS mediation analysis (section 3.3.3), that the SES measures do not have a direct effect on the outcomes but are mediated by more sensitive environmental measures. Finally, 8 – 16 variables are selected using the AIC criterion (results found in Appendix D, Figure 89) and the variables that have the highest contribution to the regression model are those selected using the BIC criterion.

	WJ.Reading	WJ.Maths	BRIEF	SDQ
Equivalised.income	-1.69	0.35	-0.63	-0.26
Parent.education.av	1.05	0.39	1.18	0.13
Parent.occupation.av	0.74	1.3	-0.12	1.2
Subjective.SES	-1.13	-2.87	-0.69	-0.19
Number.adults	0.05	-0.44	-0.55	-0.89
Number.siblings	0.35	2.58	0.93	0.96
Primary.work.hours	1.31	1.86	0.07	-2.63
RG.personal	-1.27	-2.85	0.27	-0.22
RG.acquaintance.skills	0.7	1.34	1.72	2.37
Neighbourhood.SES	1.69	0.53	1.47	0.68
Attitude.child.education	1.78	2.49	1.11	1
Child.health	-0.17	-0.95	0.91	1.2
Parent.wellbeing.and.stress	1.08	1.28	1.66	0.5
Time.with.family.and.friends	0.65	-0.08	-0.18	-0.28
Subjective.opinion.family.relationships	-2.66	-2.17	-0.79	1
Childcare	1.73	0.91	-0.69	0.83
Technology.use	-0.06	0.4	0.77	-0.44
Discipline	-2.19	0.59	-4.35	-3.97
Other.languages.spoken	-1.25	-1.03	-0.36	0.06
Reading.at.home	3.56	3.58	1.51	1.07
Attitude.neighbourhood	-1.25	0.72	-1.03	0.01
Rules	-0.32	-1.9	0.05	-0.27

Loading
2
0
-2
-4

Figure 23 OLS regression coefficients for each child outcome, averaged over imputations.
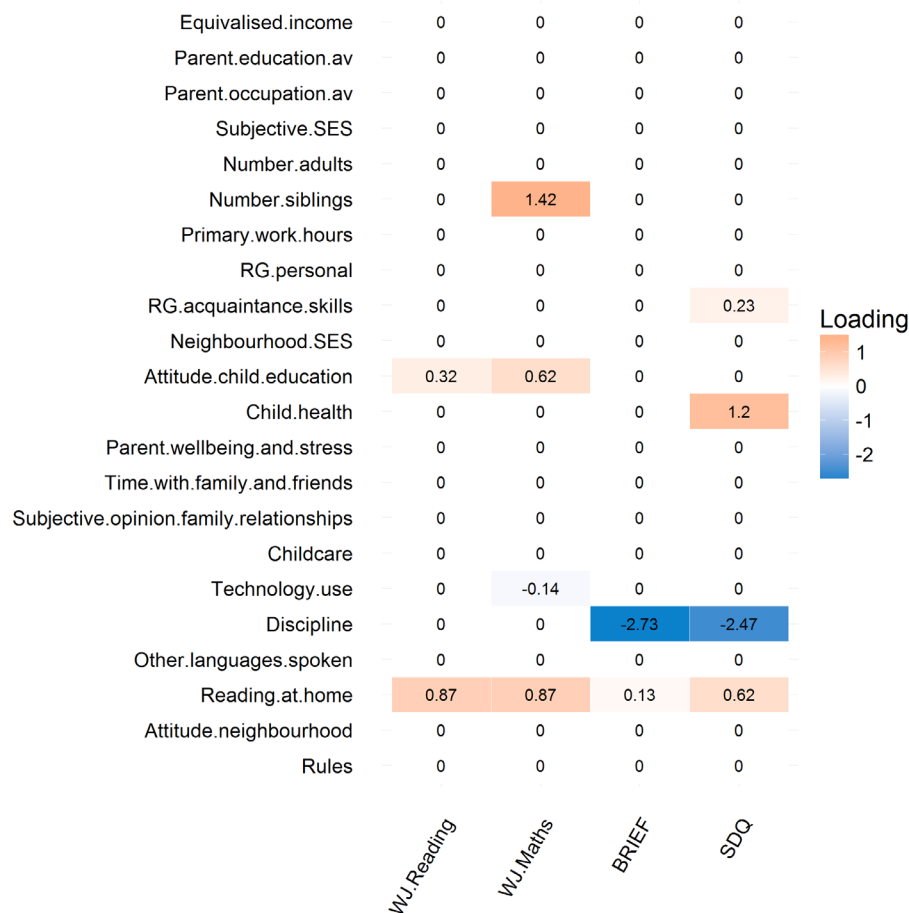
Figure 24 The selected LASSO regression coefficients for the model with the lowest BIC value

3.4 Discussion

Several studies have documented the relationship between caregiver SES and a child's academic and behaviour outcomes. Other studies have considered how specific aspects of a child's wider environment may relate to these outcomes. However, the various SES and environmental domains do not exist in a vacuum; they covary and combine to create an overall environmental profile that is related to a child's individual academic ability and behaviour outcomes. PLS methods were used to identify this profile, in addition to assessing the mediating effect of the wider environment on the SES-Outcome relationships. Following this, LASSO feature selection was used to identify a smaller subset of domains from this environmental profile that are the most predictive of each outcome.

3.4.1 The environment and academic ability

Using PLS methods, we identified a single component consisting of a wide range of domains in the child's environment that was significantly related to academic ability. Higher domain scores for reading at home, attitude to education, average caregiver occupation and education, number of siblings, child health and child care and lower technology use scores were found to collectively create an environment that related to higher academic ability. In addition, the wider environment domains also significantly mediated the relationship between child SES (caregiver education, occupation and income) and academic outcomes. The mediation model suggests that higher SES may exert a positive impact on academic ability via increased childcare, attitude to education, reading at home, child health, use of multiple languages, number of caregivers and primary caregiver skills and lower technology use, discipline and time with family and friends. These findings demonstrate the complexity of the environment-academic relationship and emphasize the importance of considering a wide range of measures in order to identify key potential targets for interventions.

Of the three SES measures, caregiver education and occupation were both found to reliably covary with the academic outcomes, a result well documented in research (e.g. (Ackerman and Brown, 2006b; Bradley and Corwyn, 2002a; Brooks-Gunn and Duncan, 1997b)). However, equivalised income did not covary with academic ability as might be expected from the income related academic attainment gap found across schools. This indicates that in this sample there are several children who have higher academic outcomes despite coming from a lower SES family and vice versa. This implies that there are other aspects in the environment that are causing the individual differences in outcomes seen, and further underlines the need to consider multiple other environmental domains.

In this analysis, reading at home and attitude to child education are two of the strongest covariates with academic ability. Attitude to education encompasses a wide variety of measures that include how the parent gives praise to the child, i.e. praise effort (2) or intelligence (1) or none (0), a child's attitude to school work and school rated by both the child and caregiver, taking part in extra-curricular activities and caregiver help with homework. Reading at home, is also predominantly based on attitudes, including measures of child and caregiver attitude to reading, the number of age-appropriate books in the home and the number of minutes a child reads or is read to a day. Therefore, both of these measures can be seen as measuring aspects of aspirations,

attitudes and behaviours (AAB's) related to education and implies that these are important domains in a child's environment for academic outcomes. This finding aligns with a number of studies, such as those summarised in a 2012 Joseph Rowntree Foundation study that considered over 170,000 pieces of evidence relating to AAB's and educational attainment. According to this study, parental involvement (often termed parent investment) in a child's education (including reading at home measures) is likely to be causally linked to education outcomes and showed the strongest association to these out of all AAB's (Stephen Gorard et al., 2012). In addition, several studies showed an association between a child's own attitude to education and academic ability, although they note that they found a lack of rigorous research into the causality of this relationship. Furthermore, in another large-scale review, specifically into attitudes towards reading, it was consistently found that this domain was one of the most predictive measures of academic ability (Education standards research team, 2012; The Reading Agency, 2015), over and above various measures of parental SES (OECD, 2002). Indeed, in this study, attitude to education and reading at home were some of the domains most strongly related to the mediating effect of the environment on the SES-Academic relationship, suggesting that these hold potential as intervention targets to improve academic outcomes in lower SES children.

The third most strongly related domain was technology use, with greater use relating to poorer academic outcomes. This measure is primarily related to the access a child has to different types of technology (including the internet, TV, mobile phone and games consoles) and the amount of time they spend using them. Note that the majority of the measures used most likely relate to technology use for entertainment. This somewhat surprising result is difficult to interpret as the association may be due to a number of factors, particularly as this domain has a relatively strong negative correlation with several other domains. For example, we might hypothesize that the greater the amount of time a child spends using technology, the less time they spend doing 'educational' activities such as reading or homework, interacting with their family and being active (Harris et al., 2017). This may be demonstrated by the negative Pearson correlation between technology use and reading at home, attitude to education and child health. There is also a strong negative correlation between technology use and caregiver education and occupation, suggesting technology use differs depending on SES. This is further demonstrated by the fact that it is found to be a key mediator of the SES-Academic relationship. This is supported by a recent large-scale study that found that

there is a socioeconomic related 'digital divide' in how children use technology (Harris et al., 2017). Specifically, children from lower SES neighbourhoods were more exposed to TV, electronic games, mobile phones, and non-academic computer activities at home and less exposed to school computers, reading, playing musical instruments, and vigorous physical activity than children from higher SES neighbourhoods (Harris et al., 2017).

The relationship found between the number of siblings and the academic component appears to be entirely driven by the relationship to maths ability only (Pearson correlations r= 0.27). It's interesting to note that, in general, this measure is inversely related to the other environmental measures, particularly income and subjective SES. The relationship with maths is in contrast to the relatively stable inverse relationship found between number of siblings and academic outcomes: higher number of siblings typically relates to lower outcomes (Downey, 1995; Gregg et al., 2007; Karwath et al., 2014). It is hypothesized that this is due to a dilution effect: that parents have fewer physical, social and emotional resources to give to each additional child (Downey, 2001, 1995). However, the number of siblings is also significantly related to the position of a child in their sibling order (in this sample, r=0.52) which has been found to show a more mixed relationship to academic ability. There is evidence, for example, that older siblings can contribute positively towards a younger child's development outcomes (Dai and Heckman, 2013), and it is possible that the effect seen here may be due to larger families being more likely to have older siblings.

The positive relationship between child health and academic outcomes has consistently been found across the literature (Eide et al., 2010). The two variables that most strongly load onto this factor are subjective ratings of the child's current and lifetime health. Other variables include the inverse of the number of days off from school, nutrition, sleep, the age of the parents at child birth, and to a lesser extent whether the mother smoked in pregnancy, birth weight and birth term. An unhealthy child is unable to engage fully with educational opportunities at school and home, which may impair their learning or cause them to miss activities. This makes it likely that their academic outcomes will be negatively impacted, as seen here. This measure is also correlated with the SES measures and is found to load reliably onto the environmental mediator of the SES-Academic relationship. SES is typically found to be strongly related to health and health-related behaviours (Bradley and Corwyn, 2002b; Pampel et al., 2010). For example, in a decomposition analysis by Gregg et al. of childhood factors associated

with the income related gradient in key stage 1 attainments, child health was found to account for 10.8% of the gradient (Gregg et al., 2007). These findings suggest that child health is both fundamental for academic ability and a potential mechanism by which SES impacts academic attainment.

Care must be taken when interpreting the relationship between childcare and academic ability as only two measures survived the skewness exclusion criteria for this study. The remaining measures related to the amount of child care the child received before school and current childcare. However, the positive association with academic ability found in this study is consistently found in research (although this depends on the quality of childcare (Bradley and Vandell, 2007; Li et al., 2013)). Childcare provides benefits such as cognitive stimulation and interaction with both adults and other children (Duncan et al., 2014; Howes, 1988). In addition, it was found to be the domain that had the strongest positive association with the mediating effect of the environment on the SES-Academic relationship, driven largely by the fact that higher SES families were more likely to have used greater amounts of childcare.

Finally, use of multiple languages, number of caregivers, primary caregiver skills, discipline and time with family and friends were all found to reliably load onto the environmental latent variable that mediated the relationship between SES and Academic attainment. However, whilst these form part of the overall picture in which children from different SES levels experience different environments, none of these domains relate strongly to the academic outcomes (according to the Pearson correlation or two-block PLS). Instead, the outer weight for each of these domains is driven by their relatively strong association with standard SES measures. This illustrates the need for caution in interpreting the importance of each variable to the mediation effect as each domain loading is the sum covariance between that domain and each of the latent variables for the other datasets.

LASSO was used to select a smaller number of environmental domains most predictive of each academic outcome. From the set of 22 environmental domains, LASSO selected two domains for reading and four for maths ability. Reading at home and attitude to education were selected for both outcomes in addition to number of siblings and negative technology use for the prediction of maths ability. This both supports some of the PLS findings that these are some of the aspects of a child's environment most related to their academic outcomes. Further, these findings indicate which variables

might be most important to measure, for example, to identify children most at risk of poor academic development related to their environment.

3.4.2 The environment and child behaviour

We also used PLS to identify a single environmental component that is significantly related to child behaviour, as measured by the BRIEF and SDQ questionnaires. Lower discipline and technology use and greater child health, reading at home, and, to a lesser extent, subjective opinion of family relationship quality, social resources (both primary caregiver and acquaintance skills), attitude to child education, caregiver education, parent wellbeing and neighbourhood SES were found to reliably covary with the behaviour component. The wider environment was also found to mediate the relationship between the SES measures and the behaviour outcomes. This mediation model suggests that higher SES may result in better behaviour outcomes via a number of domains in their environment: reduced discipline and technology use, and increased child health, primary caregiver personal skills, use of other languages, attitude to child education, reading at home, childcare, neighbourhood SES, subjective SES, acquaintance skills in the resource generator and number of caregivers. As for the academic outcomes, these PLS models demonstrate that the relationships between the environment and behaviour outcomes are highly complex and may involve several different (often related) aspects of a child's environment. Note that of the SES measures, only caregiver education was found to reliably covary with the behaviour outcomes in this sample. This implies that there are other aspects in the environment that are causing the individual differences in outcomes seen, and further underlines the need to consider multiple other environmental domains.

Discipline was found to be the domain that most strongly covaried with the behaviour outcomes. This measure consisted of 5 Likert-style questions about how often the parent used various types of discipline including time-outs, grounding, hitting, shouting or smacking. The interpretation of this measure is a somewhat complex mixture of harsher and less harsh discipline methods. A number of papers have found a relationship between higher frequency of harsh discipline techniques and poorer behaviour outcomes (Bailey et al., 2009). For example, Hecker et al. used a structural equation model analysis to show that harsh discipline is closely linked to internalising behaviour problems (measured using the SDQ) and that this in turn was associated with lower cognition and school performance (Hecker et al., 2016). Skeen et al. demonstrated that

harsh physical and psychological discipline were both related to externalising and internalising behaviours (from the SDQ) (Skeen et al., 2016). Bradley et al. found that smacking is significantly related to scores on the Behavioural Problems Index (BPI) (Bradley et al., 2001) and data from a large cohort in the Growing Up in Scotland study indicated that children who had abnormal SDQ conduct problem scores were more likely to have parents that agreed it was sometimes necessary to use smacking (Wilson et al., 2012). In addition, Gregg et al. found that frequent smacking explained 6.4% of the gradient in behaviour problems associated with income differences (Gregg et al., 2007). However, note that the domain used in the current study consists of a mixture of both harsher and less harsh discipline techniques as most of the measures correlated strongly with each other. In addition, the PLS methods identify domains that were simply associated with behaviour, rather than indicating that these domains cause behaviour outcomes. Therefore, as is noted in the Growing Up in Scotland paper (Wilson et al., 2012), it is not clear whether the discipline behaviours are in response to poor behaviour rather than a precursor of it and care must be taken as we try to interpret this relationship.

The second strongest relationship was between child health and behaviour outcomes. This is also support for this relationship in the literature. For example, the Growing Up in Scotland study also found that abnormal SDQ scores were more likely in children with poor general health (Wilson et al., 2012) and Patalay et al. found that chronic illness was associated with SDQ scores (Patalay and Fitzsimons, 2016). Gregg et al. found that child health factors explained a substantial proportion (32.8%) of the gradient in child behaviour outcomes related to income (Gregg et al., 2007). Aarons et al., found that health problem incidence was associated with mood and disruptive behaviour disorders measured using the Diagnostic Interview Schedule for Children (DISC) (Aarons et al., 2008). A study investigating health and behavioural problems in adolescents and adults with autism spectrum disorders found that prior physical health predicted subsequent behaviour problems at later time points (Kring et al., 2010). Note that in addition to child health, parent wellbeing was also found to load reliably onto the environmental component. This is made up of questions relating to low stress and the support the caregiver receives. The association between positive parent wellbeing (particularly low stress levels) and better child behaviour is also found across the literature (Baker et al., 2003; Briggs-Gowan et al., 2001; Donenberg and Baker, 1993; Neece et al., 2012; Wilson et al., 2012). Note that many of the studies referenced show

evidence for a bi-directional relationship; worse behaviour leads to increased parent stress and increased stress leads to poorer child behaviour.

There is also evidence for the finding that high technology use, particularly in terms of television viewing habits, is associated with behavioural problems in childhood. For example, greater 'screen time' has been found to relate to a variety of behavioural outcomes including aggression (Hastings et al., 2009; Mistry et al., 2007; Singer et al., 1998), anxiety and depression (Kremer et al., 2014; Singer et al., 2004), lower pro-social behaviour r (Cheng et al., 2010; Mistry et al., 2007) and attentional problems (Christakis et al., 2004; Hastings et al., 2009; Mistry et al., 2007; Swing et al., 2010) and overall conduct problems (Parkes et al., 2013). Note that these results are often highly dependent on the content of screen time, with a general trend that leisure-related screen time relates negatively to behaviour problems (particularly when the content is violent) and educational screen time relating to more positive behaviour outcomes (Hastings et al., 2009).

Reading at home and attitude to education were both identified in the PLS model as reliable covariates of child behaviour. In addition, the opinion of family relationships domains was also found to positively covary with the behaviour outcomes. This measure consists of subjective questions including how well the caregiver listens to the child and how often they discuss with their child about their day and both the child's and caregiver opinion of the quality of various family relationships. Each of these three domains contain a number of questions relating to parent-child involvement and interaction and there is a particular emphasis on the importance of this in the literature of child behaviour. For example, in a large scale literature review of family factors relating to conduct problems, parent-child involvement was one of the most powerful predictors of juvenile conduct problems (Loeber and Stouthamer-Loeber, 1986). Other studies have found positive associations between parent involvement and positive parenting (e.g. methods of praise for children) and behaviour outcomes (Adams, 2015; Dadds et al., 2003; El Nokali et al., 2010; E. M. Kim et al., 2013). Reading at home and attitude to education also contain items relating to the child's academic activities (e.g. reading for pleasure) and the child's attitudes to education. There is also evidence of this association in the literature. Gregg et al. found that the home learning environment explained 10.4% of the gradient in behaviour outcomes related to different income levels (Gregg et al., 2007). Bradely et al. found that the learning stimulation score in the

Home Observation for the Measurement of the Environment (HOME) scale is significantly negatively related to behavioural problems (Bradley et al., 2001).

Both the personal and acquaintance scales from the Resource Generator were also found to reliably covary with the behaviour outcomes. These relate to social resources in terms of primary caregiver skills and skills accessible via caregiver acquaintances respectively. To our knowledge, the association between caregiver social resources and child behaviour has not been studied before. A possible explanation for the relationship might be that as social resources increase, they overcome any lack of 'physical' resources due to low income that the family may be experiencing. Alternatively, having greater social resources may be indirectly related to behaviour. For example, higher social resources may indicate that a caregiver has better relationship and interaction skills, a greater support network or has better wellbeing and mental health which in turn may be associated with better child behaviour outcomes (Kobayashi et al., 2013; Webber and Huxley, 2007).

Neighbourhood SES had weak positive relationship to the behaviour latent variable. The neighbourhood SES domain is made up of the different indices of deprivation, used by the government to identify neighbourhood level deprivation issues. This has been found to relate to behaviour problems in other studies (Aneshensel and Sucoff, 1996; Ingoldsby and Shaw, 2002; Leventhal and Brooks-Gunn, 2000; Sampson et al., 2002), even after controlling for parental SES (Schneiders et al., 2003).

Finally, use of multiple languages, childcare, subjective SES and number of caregivers were all found to load reliably onto the environmental latent variable mediating the relationship between SES and behaviour. These were not found to be reliably associated to behaviour and the correlations between these domains and the BRIEF and SDQ scores were very low (between r = -0.07 to 0.1). Therefore, as for the academic outcomes, the outer weights can be attributed to the relatively strong association between these variables and the standard SES measures and they are not likely to be good candidates for interventions aimed at improving behaviour outcomes.

LASSO was used to select a smaller number of environmental domains most predictive of each behaviour outcome. From the set of 22 environmental domains, LASSO selected two domains for the BRIEF and four for the SDQ. The strongest predictor by far was the discipline domain, where more discipline use by parents was predictive of poorer behaviour outcomes. The LASSO results suggest that this is an excellent

predictor of child behaviour, whether or not discipline is a response to or precursor of poor behaviour as described above. This may be of interest to practitioners as an alternative measure in contrast to longer questionnaires such as the BRIEF or SDQ as it consists of only 5 Likert- style questions. In addition, positive reading at home was selected for both behaviour outcomes. It is interesting that this appears to be such an important predictor for each of the academic and behaviour outcomes. Finally, positive child health and acquaintance skills (Resource generator questionnaire) were selected for the prediction of the SDQ scores, indicating that these are also important predictors to include when considering child behaviour outcomes.

In summary, PLS and LASSO offer two complementary methods for investigating the complex relationships between the domains of a child's environment and their academic and behaviour outcomes. PLS summarised the overall environmental profiles that most strongly covary with individual outcomes, in addition to offering insight into how the wider environment may mediate the well-studied relationship between SES and these outcomes. A number of environmental domains were found to reliably covary with the child outcomes and the SES-outcome relationship. These help to highlight potential environmental risk factors in children and identify key targets for future interventions that support these children. In turn, LASSO identified a small number of domains amongst this complex set that most strongly predict unique variance in each outcome. Identifying these in the presence of high collinearity from the set of environmental domains aids practitioners and researchers to avoid measuring redundant domains whose covariance with the outcome is better explained by other measures.

4 ENVIRONMENT AND COGNITION

4.1 Introduction

Cognitive skills refer to our ability to perform mental activities such as processing information, reasoning, problem solving, learning and memory. These skills are key developmental markers of interest to practitioners as they underpin many of our behaviours and abilities. For example, strong associations have been found between cognition and academic attainment (Alloway and Alloway, 2010; Colom et al., 2007; Deary et al., 2007; Rohde and Thompson, 2007; St Clair-Thompson and Gathercole, 2006; Welsh et al., 2010), internalising and externalising behaviour (Blanken et al., 2017; Pennington and Ozonoff, 1996) and a variety of positive lifetime outcomes (Hofer and Clouston, 2014) including higher education attainment (McClelland et al., 2013), career outcomes (Heckman et al., 2006) and health and wellbeing (Heckman et al., 2006; Llewellyn et al., 2008; Richards et al., 2010).

In addition, many cognitive skills have a prolonged developmental period that extends throughout childhood and into adulthood (Best and Miller, 2010; Casey et al., 2000; Keating, 2013; Steinberg, 2005), making them particularly susceptible to differences in child environment. For example, experience dependent synapse formation, pruning and myelination are protracted throughout the early childhood years in areas of the brain that strongly relate to language and continue well into the late teenage years in the prefrontal cortex, critical for many higher cognitive abilities (Casey et al., 2000, 1997; Giedd et al., 1999; Huttenlocher and Dabholkar, 1997; Teffer and Semendeferi, 2012). As such, cognitive skills are likely to be strongly associated to aspects of a child's environment and are particularly interesting candidates for mediating the relationship between a child's environment and their academic and behaviour outcomes.

Research has identified several strong associations between standard SES measures and cognitive skills, particularly for language-related skills such as vocabulary and phonological processing and higher cognitive abilities such as working memory and

fluid intelligence (Farah et al., 2006; Noble et al., 2007, 2005). However, little has been done to build on these simplistic relationships by investigating which aspects of a child's environment are most closely associated with cognition or exploring the mediating role of different cognitive skills for key childhood outcomes like attainment or behaviour.

In order to address this, we used PLS methods to identify the profile of environmental variables that most strongly relate to a child's cognitive skills. Working memory, vocabulary, matrix reasoning (fluid intelligence) and phonological processing were selected as key cognitive abilities typically found to be particularly susceptible to SES effects (Noble et al., 2007, 2005). The relationship between these skills and multiple aspects of a child's environment were explored to answer three questions. First, what are the key aspects of a child's environment most strongly associated with their cognitive profile? Second, does the wider environment mediate the relationship between the standard SES measures and cognition? Finally, does cognition act as a potential mediating mechanism underlying the relationship between environment and academic and behaviour outcomes?

4.2 Methods

4.2.1 Participants and tasks

This analysis used the same sample as the previous analysis: 94 children (45 males, mean age 9:10 and age range 6:11-12:9). In addition to the 21 environmental domains and the academic ability and behaviour outcome scores, age standardised cognitive skills were measured using a battery of 5-10-minute tasks. Verbal and Fluid IQ were measured using the Verbal and Matrix Reasoning subtests from the Wechsler Abbreviated Scale of Intelligence II (WASI-II) (Wechsler, 2011). Phonological processing was assessed using the Naming Speed subtest from the Phonological Assessment Battery (PhAB) (Frederickson and Reason, 1997). Short-term and working memory (STM and WM) was assessed using four subtests of the Automated Working Memory Assessment (AWMA): Digit Recall, Backwards Digit Recall, Dot Matrix and Mr X (Alloway, 2007). These are designed to tap verbal STM, verbal WM, visuo-spatial STM and visuo-spatial WM respectively. A factor analysis was conducted on the four working memory subtests, and the first factor was selected based on the factor

eigenvalues, the position of the elbow in the scree plot and parallel analysis (detail in chapter 2).

4.2.2 PLS between the environment and cognition

The Pearson correlation between each of the cognitive scores and the environmental domains and outcome measures was calculated. PLS was then used to identify the profile of environmental factors that most strongly covaries with the set of cognitive skills. Each variable was standardised and the PLS analysis was conducted as in section 3.2.2. with the environmental variables and cognitive skills used as dataset X and Y respectively. Rubin's rules were used to pool the PLS results across imputation and the significance and reliability of the PLS model were assessed using permutation (N permutations = 1000) and bootstrap (N bootstrap samples = 500) analyses.

4.2.3 PLS mediation

We applied 3-block Canonical PLS to the standardised datasets as detailed in section 3.2.3 to investigate mediation effects. First, the mediating effect of the wider environment on the SES-cognition relationship was investigated using the three standard SES measures, equivalised income, caregiver education and occupation as dataset X, the other environmental domains as dataset M and the four cognitive skill scores were used as dataset Y. Note that, as in section 3.2.3, Primary caregiver's hours work per week was excluded from the wider environment dataset as it is very closely related to equivalised income.

Secondly, in order to investigate whether cognitive skills mediate the relationship between a child's environment and their academic and behaviour outcomes, the full set of environmental domains were used as dataset X, the four cognitive skill scores were used as dataset M and either the two academic variables or two behaviour variables were used as dataset Y.

4.3 Results

4.3.1 Pearson correlation results

The four cognitive skills were moderately to highly correlated with each other (r = 0.3-0.55). The relationships between the cognitive skills and the different environmental domains were mixed (Figure 25). The majority demonstrated a positive relationship, but there were also several correlation coefficients that were either very small or negative.

In general, the patterns of correlation were similar for the working memory and matrix reasoning scores, reflecting the relatively high correlation between these two cognitive skills ($r = 0.55$). Phonological processing appeared to be weakly related to the majority of environmental domains. Note that of the standard SES measures, equivalised income did not relate to any of the cognitive skills in this sample, whereas caregiver education and occupation levels did.

There was a relatively strong positive correlation between all of the cognitive skills and the academic outcomes, shown in Figure 26 ($r = 0.34$-0.61). The relationships were also positive for the BRIEF, but the strength of these correlations was much lower ($r = 0.09 – 0.36$). Only matrix reasoning appeared to correlate with SDQ scores. Phonological processing was not correlated with each of the behaviour outcomes.

	Working.Memory	Matrix.Reasoning	Vocabulary	Phonological.processing
Equivalised.income	-0.09	0.04	0.01	-0.11
Caregiver.education.av	0.23	0.3	0.17	0.11
Caregiver.occupation.av	0.12	0.28	0.29	0.08
Subjective.SES	-0.13	-0.04	0.17	0.03
Number.caregivers	0.06	0.03	0.09	0.14
Number.siblings	0.26	0.19	0.11	0.22
Primary.caregiver.hours.work	0.04	0.15	0.22	0.08
Primary.caregiver.skills	-0.06	0.03	0.09	0.01
Acquaintance.skills	0.09	-0.01	0.31	0.07
Neighbourhood.SES	-0.06	0.04	0.13	0.03
Attitude.child.education	0.11	0.12	0.16	0.16
Child.health	0.13	0.17	0.22	0
Primary.caregiver.wellbeing	-0.1	-0.18	0.08	-0.1
Time.with.family.and.friends	-0.09	-0.21	0.09	-0.03
Opinion.family.relationships	-0.09	-0.06	0.11	-0.23
Childcare	0.1	0.15	0.1	-0.03
Technology.use	-0.28	-0.35	-0.04	-0.09
Discipline	-0.19	-0.23	-0.19	0.05
Other.language.use	0.06	0.13	-0.1	-0.03
Reading.at.home	0.21	0.19	0.3	0.13
Attitude.to.neighbourhood	0	-0.15	-0.03	0.02
Rules.and.chores	-0.18	-0.12	0.21	0.1

Figure 25 The Pearson correlation between each environmental domain score and the cognitive ability

Figure 26 The Pearson correlation between the cognitive ability and the academic and behaviour outcomes

4.3.2 PLS: What is the environmental profile that most strongly relates to a child's cognitive profile?

Two-block PLS was conducted using the environmental and cognitive datasets. The first pair of latent variables significantly explained the covariance between the environmental domains and the cognitive skills (p = 0.008, explained variance: 15% environment and 55% cognitive scores, average correlation between environment and cognitive latent variable = 0.50). Working memory, matrix reasoning and vocabulary each loaded strongly onto the first component (Figure 27). Caregiver education and occupation, reading at home, number of siblings, child health and attitude to child education show a strong to medium positive relationship with this cognitive profile. Technology use and discipline are inversely related to the cognitive skills.
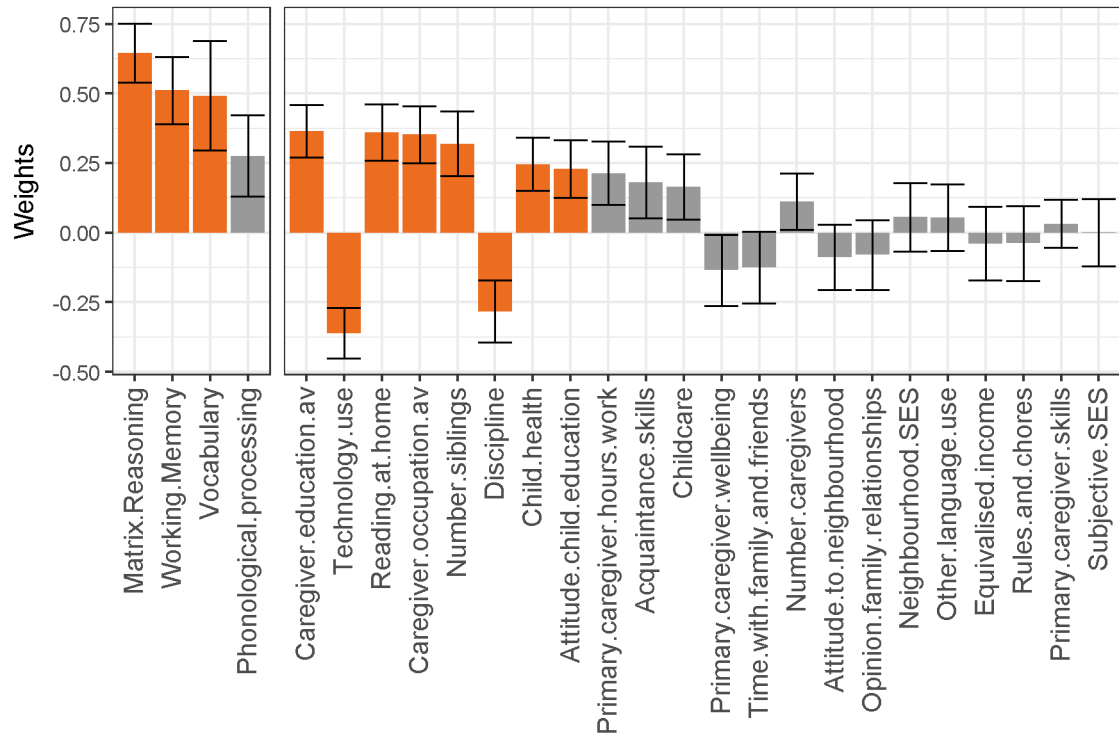
Figure 27 The environment and cognitive ability: outer weights for the first latent variables, averaged across imputations. Error bars are the standard errors obtained from the bootstrap samples, pooled across imputations using Rubin's rules. Reliable loadings (according to the critical t-value) are coloured orange.

The second pair of latent variables was also significant ($p = 0.01$, explained variance: 13% environment and 15% cognitive scores, average correlation between environment and cognitive latent variable = 0.53). Illustrated in Figure 28, this component captures the positive covariance between the vocabulary skills and rules and chores, acquaintance skills, subjective SES and technology use, after the first component is removed from the data. This implies that vocabulary skill is related to unique variance in a child's environment. However, whilst it is interesting to identify the components that explain significant amounts of covariance amongst the datasets, it is difficult to interpret what this means as the second component environment weights are proportional to the covariance explained by each variable after the variance explained by the first component has been removed. Caution must therefore be taken in applying these higher order results to practical application such as design of interventions.
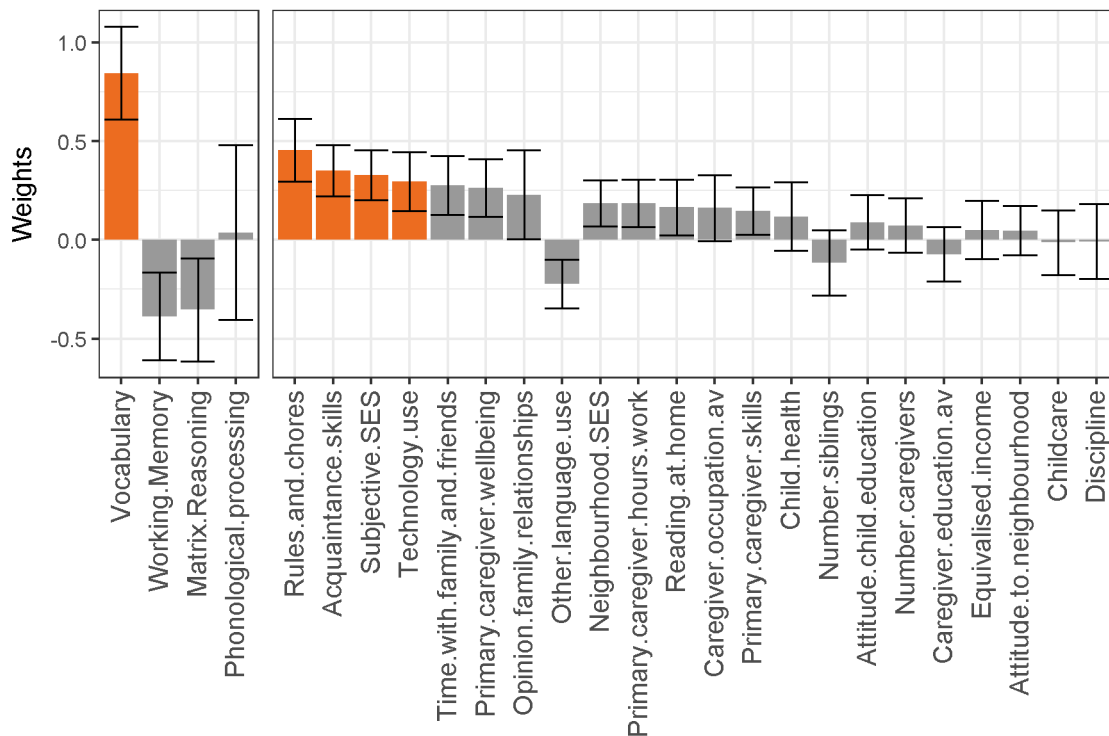
Figure 28 The environment and cognitive ability: outer weights for the second latent variables, averaged across imputations

4.3.3 PLS mediation results

4.3.3.1 Does the wider environment mediate the relationship between standard SES measures and cognition?

The wider environment significantly mediated the relationship between SES and a child's cognition for the first three sets of latent variables. As we are focussing on identifying the key targets for potential interventions and it is difficult to interpret the associations in higher components, only the first component will be considered here. The environment domains fully mediated the SES-Cognition relationship (indirect path coefficient ab=0.26±0.09, 95% CI [0.16,0.49]). The path model for the first component can be seen in Figure 29 and the variable weights can be seen in Figure 30.
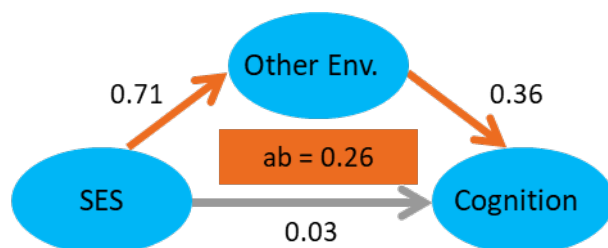
Figure 29 PLS mediation analysis between SES, the wider environment and cognition: the path model
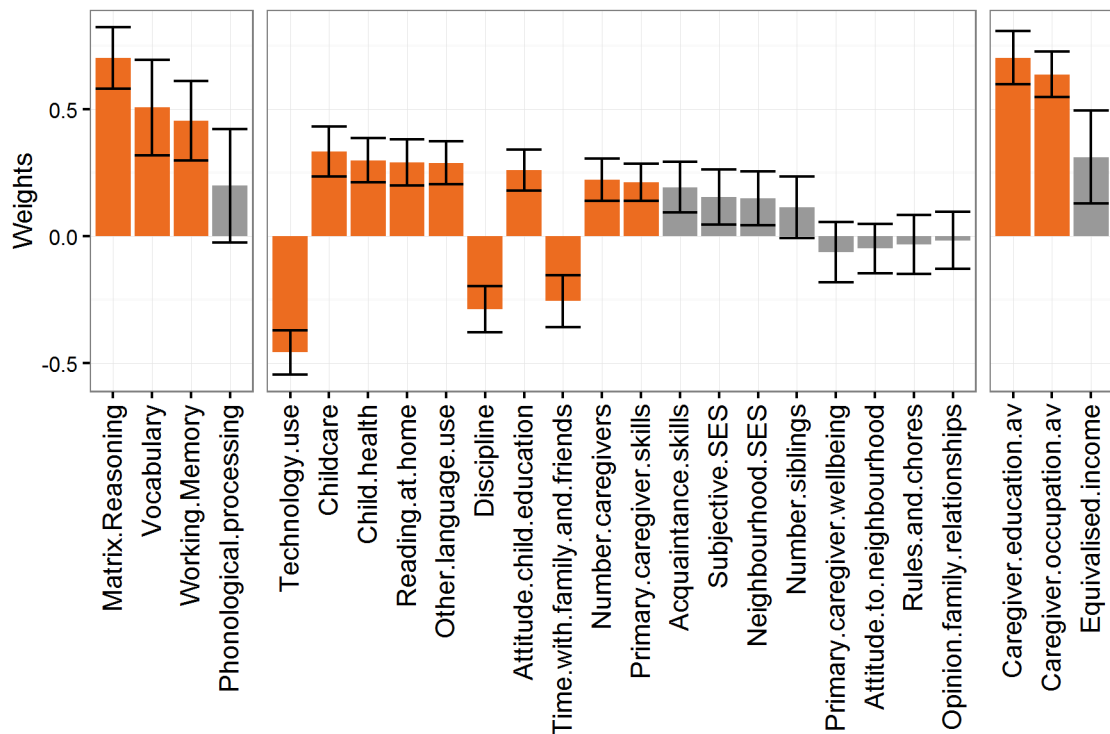
Figure 30 PLS mediation analysis between SES, the wider environment and cognition: outer weights for the first latent variables, averaged across imputations

The pattern of loadings was very similar for the pairwise model between all of the environmental measures and the cognition. However, there were a few additional environmental measures that were reliably associated with the mediation effect of the wider environment on the SES-Cognition relationship: childcare, other language use, inverse time with family and friends, number of caregivers and primary caregiver skills.

4.3.3.2 Do cognitive skills mediate the relationship between a child's environment and academic ability?

The mediation analyses above and in section 3.3.3 indicated that the standard SES measures did not explain unique variance in either the cognitive, academic or behaviour outcomes and so the following mediation analyses considered all environmental domains together as one dataset.

The mediating effect of cognitive skill on the environment-academic relationship was significant for the set of latent variables in the first component (indirect path coefficient $ab=0.31\pm0.05$, 95% CI [0.25,0.45]). The path model can be seen in Figure 31 and the variable weights can be seen in Figure 32. The second component was not found to be significant.
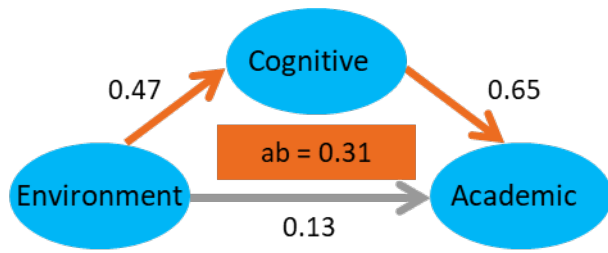
Figure 31 PLS mediation analysis between the environment, cognition and academic ability: the path model

Both reading and maths loaded onto the component. A number of environmental domains loaded onto the component in order of magnitude: reading at home, inverse technology use, caregiver education and occupation, attitude to child education, number of siblings, child health and inverse discipline. Finally, each of the cognitive skills reliably related to this mediating effect. The magnitude of each of the variable outer weights relate to the sum covariance between each variable in one dataset and the latent variable of each of the other datasets and so the weights indicate the importance of each variable for explaining the three-way relationship.
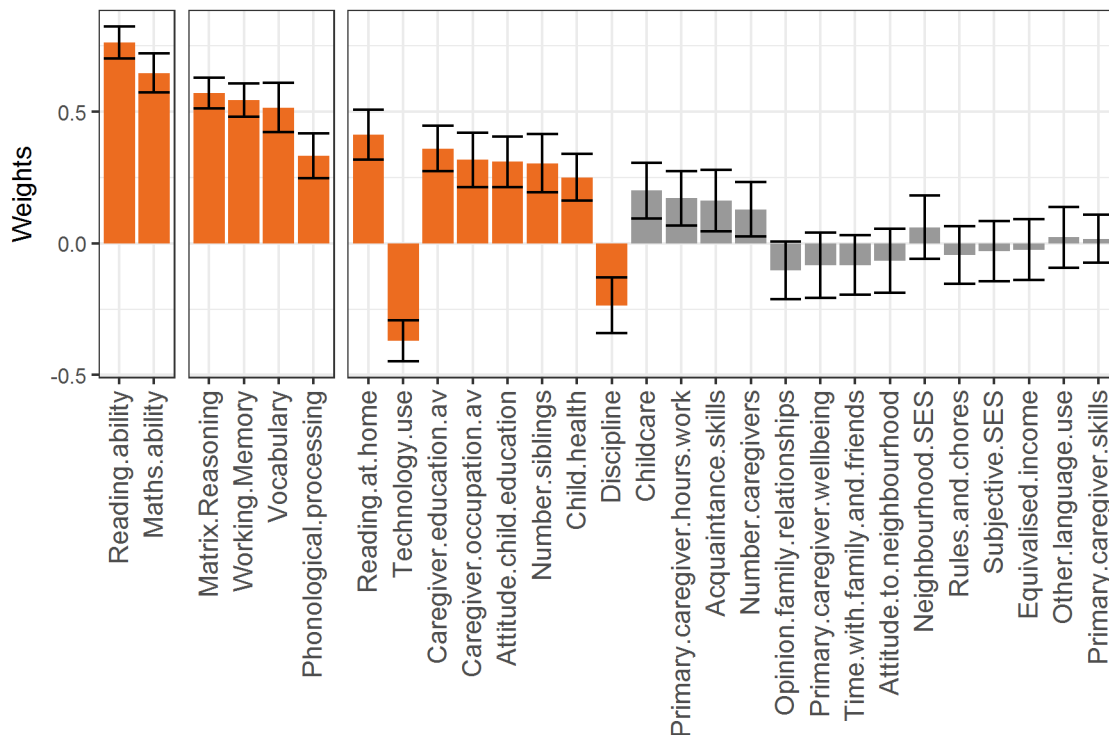
Figure 32 PLS mediation analysis between the environment, cognition and academic ability: outer weights for the first latent variables, averaged across imputations

In order to visualise this relationship, the environment latent variable is plotted against the academic latent variable in Figure 33, demonstrating the positive relationship between the environment and academic ability. Furthermore, the participant scores are coloured according to their cognitive score based on a median split of this score. Higher academic and higher environment scores both related to higher cognitive scores and vice versa, which related to the mediation effect found.
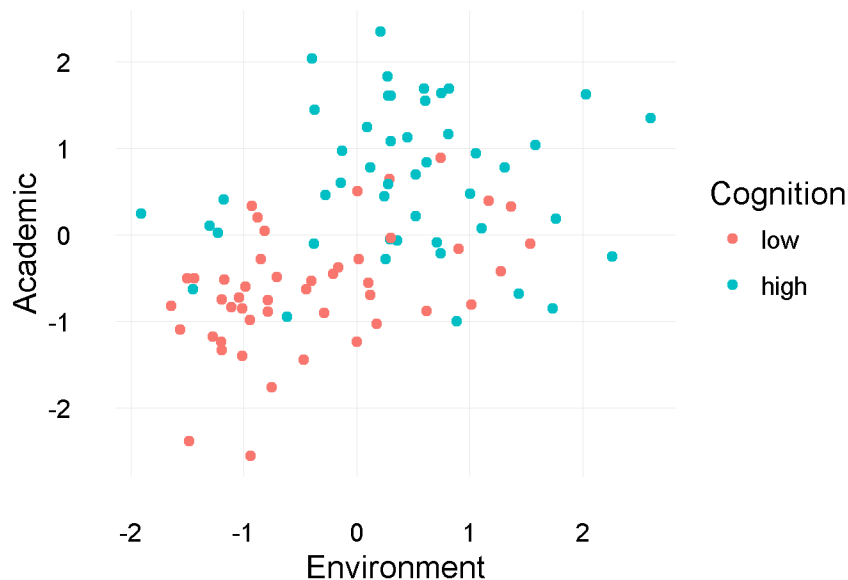
Figure 33 PLS mediation analysis between the environment, cognition and academic ability: scatter plot illustrating the relationship between the SES and academic latent variables, coloured by median split of the wider environment domain.

4.3.3.3 Do cognitive skills mediate the relationship between a child's environment and behaviour outcomes?

The mediating effect of the environment on the cognition-behaviour relationship was not significant for any of the components (First component ab=0.03±0.03, 95% CI [-0.05,0.08]). The outer weights indicated the sum covariance between each variable in a dataset and the two latent variables from the other datasets, given in Figure 35. BRIEF and SDQ are both reliably related to the cognitive and environment datasets. Inverse discipline, reading at home, child health, caregiver education, inverse technology use, caregiver occupation, attitude to child education, caregiver acquaintance skills, number of siblings and caregiver personal skills were reliably related in order of strength to the other datasets. Matrix reasoning, vocabulary and working memory each covaried reliably with the environment and child behaviour outcomes.
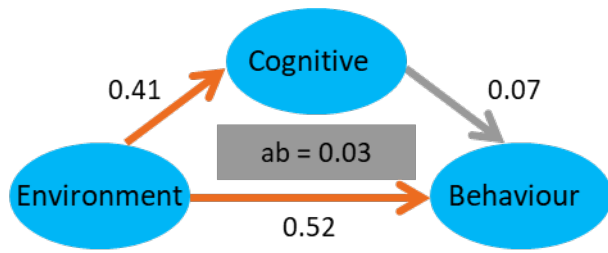
Figure 34 PLS mediation analysis between the environment, cognition and behaviour: the path model
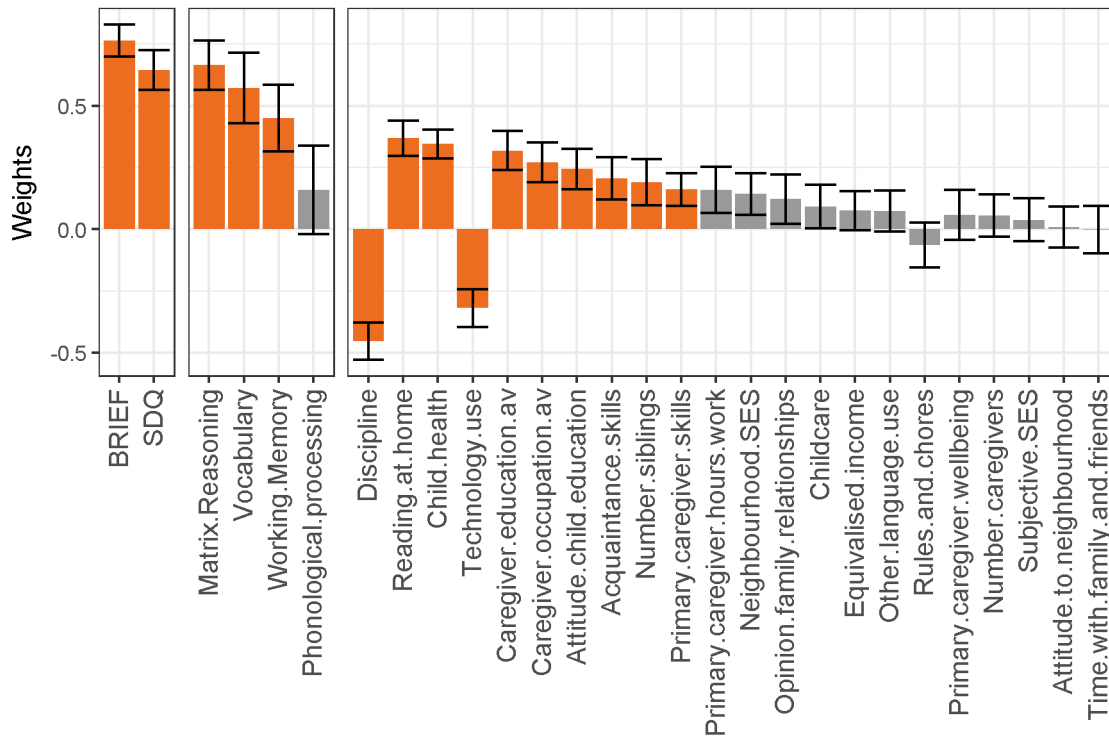
Figure 35 PLS mediation analysis between the environment, cognition and behaviour: outer weights for the first latent variables, averaged across imputations

The path analysis indicated that the environment explains a large proportion of variance in the behaviour outcomes, even when cognitive skills were held constant. Indeed, the pairwise PLS results of the Environment-Behaviour relationship in 3.3.2.2 identified a number of environmental domains that reliably covaried with behaviour, that were not found to covary with the cognitive latent variable. In particular, opinion of family relationships, primary caregiver skills, primary caregiver wellbeing, acquaintance skills and neighbourhood SES were all found to covary with behaviour and not cognition. However, there were also environmental domains found to covary with both cognition and behaviour: inverse discipline, child health, reading at home, inverse technology use, caregiver education and attitude to education. This is illustrated in Figure 36.
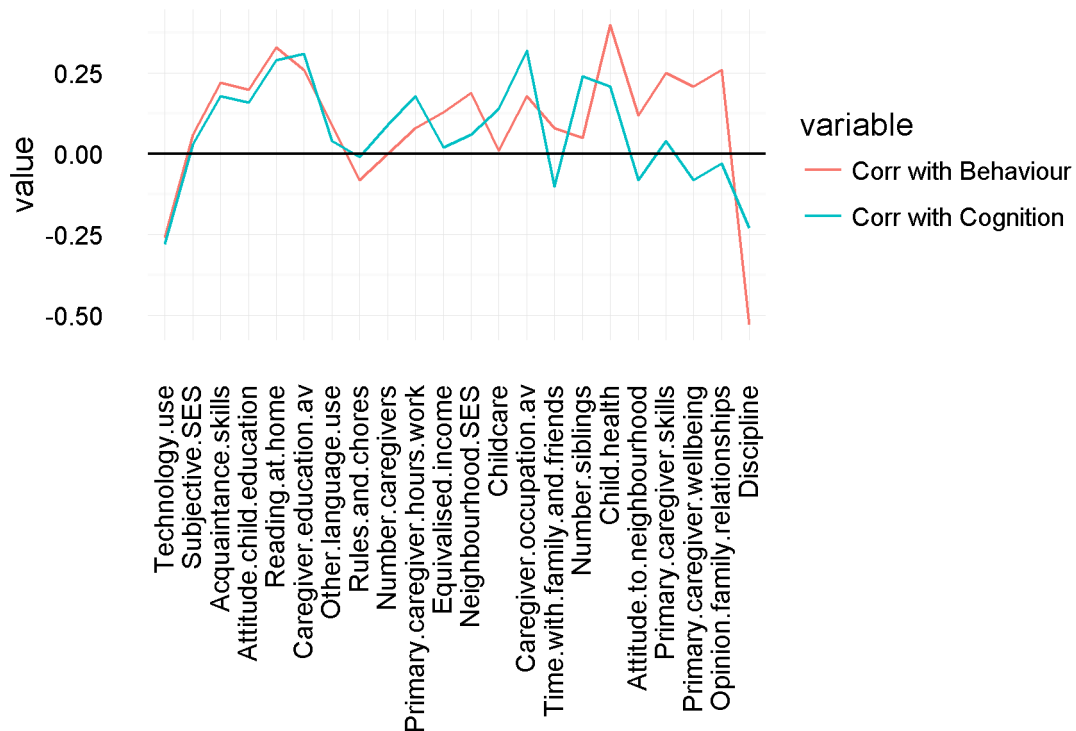
Figure 36 The correlations between each environmental domain and the cognitive and behaviour latent variables from the PLS mediation analysis, ordered by the difference in correlation coefficient.

In addition, note that the correlation between cognition and behaviour was significant (r=0.28±0.07, 95% CI [0.15,0.43]) for the first set of latent variables, but holding the environment constant eliminated this relationship. The fact that differences in the child environment fully explained the association between cognition and behaviour suggests that this association may not be causative and instead may only appear to be correlated because they both covaried with a third covariate (the environment). That is, relationships between cognition and behaviour may not be due to direct mechanisms, but because both are strongly associated with aspects of a child's environment, they at least partially overlap.

4.4 Discussion

4.4.1 The environment and cognition

This research moves beyond identifying simple pairwise relationships between standard SES measures and cognition to investigate how multiple aspects of a child's environment may be associated with different cognitive skills. The PLS results demonstrate that several environmental domains covary and combine to create an

overall environmental profile that is related to a child's set of cognitive abilities. In particular, an environment characterized by lower levels of technology use, greater enjoyment and time spent reading at home, more siblings, lower levels of discipline, good child health and positive attitude to education were associated with higher ability in working memory, matrix reasoning and vocabulary. The standard SES measures, caregiver education and occupation, were also found to reliably covary with cognitive skills. However, as for academic and behaviour outcomes, equivalised income did not covary with cognitive ability, demonstrating again that in this sample there were several children who have high cognitive skills despite coming from a lower SES family and vice versa. Indeed, individual differences in the wider environmental domains fully explained the relationship between these standard SES measures and cognition. This underlines the need for developmental studies to consider multiple aspects of the environment beyond standard measures of SES. Furthermore, it identifies potential targets for future intervention studies aimed at reducing the negative impact of growing up in a deprived environment on the development of cognitive skills.

After caregiver education, the strongest association was found for technology use: lower use of technology related to higher cognitive abilities. It was also the domain most strongly associated with the mediation of the SES-Cognition relationship. The Pearson correlations indicated that this association is primarily driven by the relationship of this domain to working memory and matrix reasoning, tasks that are typically seen to come under the umbrella of executive functions rather than to vocabulary and phonological processing. It is widely accepted across the literature that both positive and negative associations between cognition and technology exist depending on several factors including the content, context, level of active engagement, duration and the timing of technology exposure during a child's development (Bavelier et al., 2010; Schmidt and Vandewater, 2008). For example, technology use that is focussed on entertainment at home, the primary type of technology use assessed in this study, is typically more likely to be inversely related to cognitive skill, in contrast to technology use with higher educational content or technology use at school (Huber et al., 2018; Lillard and Peterson, 2011; Nathanson et al., 2014; Schmidt and Vandewater, 2008). As described in chapter 3, one of the leading theories for the mechanism of this effect is that technology use for entertainment replaces other activities more beneficial to the development of cognitive skills (Comstock and Paik, 1991), such as reading at home which has a strong positive association with cognition (Vandewater et al., 2005).

Alternatively, technology use covaries with several other strong predictors of cognitive skill and controlling for these has been shown to remove the association between technology use and cognition (Schmidt et al., 2009; Schmidt and Vandewater, 2008), suggesting that this domain may not be causative. For example, technology use may simply be a particularly strong proxy measure for a child's underlying environmental profile. However, the strength of the association suggests that technology use, particularly in the context of entertainment at home, is an important candidate for interventions and further studies into the causality of this relationship are needed.

Reading at home is the second strongest covariate of cognitive skill. In the literature, the importance of the home reading culture is particularly well established for the development of vocabulary (Byford et al., 2012; Iltus, 2006; Teale, 1981). For example, a longitudinal study found that childhood reading is associated with substantial progress in vocabulary between the ages of 10 and 16, and that this provided a greater advantage than caregiver education (Sullivan and Brown, 2013). Whilst few studies have investigated the specific relationship between reading at home and other cognitive skills, aspects of reading at home have been investigated under wider domains such as 'parent involvement' or 'home learning environment and materials'. Note that these domains often also capture measures relating to attitude to child education, also found to covary reliably with cognition in this study. For example, Bradley et al. found that the strongest correlations between IQ and environmental domains were the aspects that relate to cognitive stimulation such as availability of learning materials and parent involvement in child activities from the Home Observation for the Measurement of the Environment (HOME) scale (Bradley et al., 1989). The home learning environment accounted for 9.4% of the gradient in IQ scores across SES levels in the decomposition analysis of the ALSPAC cohort by Gregg et al. (Gregg et al., 2007). The availability of books and toys, and paternal reading and singing to the child accounted for the largest shares of this. Hughes and Devine found that the home learning environment, which included the frequency of reading at home, was predictive of child executive function and verbal ability in 5-year-olds (Hughes and Devine, 2017). Sarsour et al. applied the HOME scales to investigate mediation of the relationship between SES and child executive functions of inhibitory control, cognitive flexibility, and working memory. Enrichment (the use of family and community resources to enrich the development of the child), and family companionship (parent involvement in child activities)

significantly mediated the relationship between family SES and child inhibitory control and working memory (Sarsour et al., 2011).

We found that higher numbers of siblings related to higher cognitive abilities, with a positive correlation between each of the cognitive skills and sibling number. This is in contrast to the reverse relationship typically found in the literature (Downey, 2001), the majority of which has focussed on the inverse relationship between family size and intelligence (Holmgren et al., 2006). In one of the most famous examples, Zajonc and Markus demonstrated that larger family size was significantly inversely related to Raven's Progressive Matrices score (similar to matrix reasoning) across the entire male population of the Netherlands who turned 19 between 1963-1966 (Zajonc and Markus, 1975). A negative relationship has also been demonstrated between vocabulary and sibling number (Downey, 1995; Heiland, 2009). However, the sibling effect is often marginal (Holmgren et al., 2006) and there are examples in which no significant relationship was found for sibling number and vocabulary (Cole and Mitchell, 2000), fluid intelligence (Holmgren et al., 2006) and working memory (Holmgren et al., 2006). In addition, as for academic attainment, older siblings may positively contribute towards their younger sibling's cognitive development. For example, the quality of sibling interactions has been found to provide a protective factor that moderates the negative impact of large sibship size on language development (Prime et al., 2014). It is possible that there are additional sibling factors in our sample that protected against the negative effects of resource depletion in larger families and further research is needed to investigate the complex effects of siblings on child development. Note that number of siblings was not found to reliably relate to the mediation of the SES-Cognition relationship. This is not surprising as the Pearson correlations in section 3.3.1 indicate that number of siblings is inversely related to equivalised income and does not relate to caregiver education or occupation.

Better health is consistently related to higher cognitive ability, as found in this study. Several literature reviews of the relationships between a number of health promoting behaviours and cognition have identified both short- and long-term effects on cognition, such as for good nutrition (Sorhaindo et al., 2006) and higher levels of physical activity (Sibley and Etnier, 2003; Tomporowski et al., 2015). In addition, child health is one of the domains most strongly related to the mediation of the relationship between SES and cognition. Health is closely linked to a family's SES and is hypothesized to be one of the key mechanisms mediating the relationship between SES and cognition (Bradley

and Corwyn, 2002b). For example, Gregg et al. found that health behaviours were not only correlated with IQ, but also accounted for a significant proportion of the gradient in IQ scores due to SES levels (Gregg et al., 2007). Furthermore, health behaviours are often the target of intervention studies, establishing a causative relationship between factors. From these, improving health behaviours has been found to improve cognitive abilities in child interventions in both developed and developing countries (Engle et al., 2007; Kamijo et al., 2011; Tomporowski et al., 2015).

Discipline is found to be reliably inversely related to cognitive ability. This relationship is typically found in the literature for greater levels of harsher discipline techniques. For example, Byford et al. found that the use of threats and coercion was significantly inversely related to a child's vocabulary (Byford et al., 2012). Noble et al. reported that the frequency of physical punishment is negatively related to vocabulary and working memory (although this was not significant when other home and school environmental measures were accounted for) (Noble et al., 2007). Gregg et al. found a highly significant inverse relationship between IQ and the frequency of smacking, that accounts for 2% of the IQ gradient related to family SES (Gregg et al., 2007). However, note that in the current study, the less harsh discipline techniques such as using negative consequences are relatively strongly correlated with the harsher discipline techniques such as shouting and smacking and so the discipline domain score is a mixture of both. In further studies, it will be important to determine whether this strong inverse relationship is primarily driven by the harsher discipline styles or discipline in general.

Childcare, time with family and friends, number of caregivers, and primary caregiver skills were each reliably related to the mediation effect of the wider environment on the relationship between SES and cognition. However, these domains were not found to covary with the cognitive latent variable in the PLS model between all environmental variables and cognition. This indicates that the reliable loading onto the mediation latent variable is due to the high covariance between these domains and the SES measures, rather than covariance with cognition.

Finally, it should be noted that phonogical processing was not found to reliably relate to the environment latent variable. In general, the Pearson correlations between this and the individual environmental domains remained close to zero. A possible reason for this is that only one task, a timed task that emphasizes the speed of naming pictures, was used to obtain a measure of phonological processing ability. There are, in fact, several other aspects of phonogical processing and these may be found to be more strongly

associated with a child's environment. Indeed, SES has been consistently related with phonological skills (Eckert et al., 2001; McDowell et al., 2007; Noble et al., 2007), and further research should consider the associations between a wider battery of phonogical tasks and the multiple environment domains.

4.4.2 Cognition mediated the relationship between the environment and academic outcomes

Each of the cognitive abilities were strongly correlated with the academic outcomes. In addition, the environmental domains most strongly associated with cognition were also those that covaried most strongly with the academic outcomes. Therefore, it is not surprising that cognition was found to fully mediate the relationship between a child's environment and their academic ability. One explanation of this might be that a 'positive' environment may enable healthy development of cognitive skills, which in turn supports the child in academic learning, suggesting a potential mechanism by which the environment may impact child outcomes.

These results are correlational and do not imply a causative pathway. However, there is evidence of a causal link between cognitive skills and academic ability in several longitudinal studies. Early working memory skills in particular were found to predict growth in maths (Passolunghi et al., 2007; Raghubar et al., 2010; Swanson and Jerman, 2007) and reading ability (Cain et al., 2004; Seigneuric and Ehrlich, 2005) over time. For example, Welsh et al. found that growth in working memory and attentional control skills across the pre-kindergarten year predicted maths and reading ability at the end of the kindergarten year, after controlling for growth in pre-kindergarten literacy and numeracy skills (Welsh et al., 2010). The link between early vocabulary skills and later reading is also well studied (e.g. see the literature reviewed in (Oakhill and Cain, 2012). For example, Oakhill et al. found that vocabulary measured in year 3 of school made a significant unique contribution to reading comprehension in year 6 and Verhoeven and Van Leeuwe demonstrated a similar effect across 6 years of primary school (Verhoeven and Leeuwe, 2008). In a large longitudinal study of 70,000 British children, Deary et al. demonstrated that fluid reasoning ability at age 11 correlated very strongly ($r = 0.81$) with GCSE results summarised as a latent variable (Deary et al., 2007). However, note that academic ability at age 11 was not controlled for.

Furthermore, a few studies have considered the mediating impact of cognition on the relationship between measures of the environment and academic ability. For example,

Lawson and Farah found that executive function measured at the first time point (primarily consisting of measures of working memory) significantly mediated the relationship between SES and change in maths ability over two years (Lawson and Farah, 2017). They also found that there was some mediation of the SES-reading relationship, but this was not significant. Nesbitt et al. found that kindergarten executive functioning skills mediated the relationship between SES and grade 1 maths and literacy (Nesbitt et al., 2013). In an older group of 8-11 year olds, Crook and Evans found that planning skills (an aspect of executive functions) partially mediated the relationship between income to needs ratio and maths ability two years later (Crook and Evans, 2014). In a cross sectional study, Dilworth-Bart et al. found that executive function mediated the relationship between SES and maths readiness at the start of school (Dilworth-Bart, 2012). Finally, Alves et al. found a strong indirect path coefficient between SES and academic ability via IQ in 4-10 year olds (Alves et al., 2017). Taken together, these studies provide further evidence that cognitive skills might facilitate a mechanistic pathway between a child's environment and their academic outcomes, in line with our results.

The majority of research to date has focussed on the mediation of the relationship between standard measures of SES and academic ability. Only two studies we know of investigated the mediating effect of cognitive skills on the relationship between wider environmental domains and academic outcomes. Using the NICHD Study of Early Childcare, sustained attention and impulsivity mediated the relationship between a composite score consisting of the total home environment quality measured using the HOME questionnaire, maternal sensitivity and maternal cognitive stimulation and both reading and maths achievement (NICHD Early Child Care Research Network, 2003) at 54 months old. Secondly, Devine et al. found that executive functions mediated the relationships between parenting behaviours such as academic scaffolding and negative parent-child interactions at age 4 and child academic ability 1 year later (Devine et al., 2016).

Our study builds on this work by demonstrating mediation of the relationship between a set of multiple environmental domains and academic outcomes. In addition, most research has been conducted with children in preschool or the first few years of schooling. This study provides evidence that cognitive skills may provide a potential mechanism for the relationship between the environment and academic abilities in older children. Furthermore, studies have either focussed on single cognitive skills or enter

different cognitive skills (e.g. working memory and vocabulary) as separate latent variables in a path analysis, thus controlling for the other cognitive skills in each indirect path. However, cognitive skills covary strongly and it is also useful to consider their relationship to the mediation effect without partialling out this covariance. By using PLS methods, we have identified the cognitive profile that is most strongly associated with the mediation of the relationship between the environment and academic ability.

4.4.3 Cognition did not mediate the relationship between the environment and behaviour outcomes

Cognition was not found to mediate the relationship between the environment and behaviour. There are two key results from this. The first is that the environment explains a large amount of variance in a child's behaviour outcomes, even when their cognition is held constant, emphasising the importance of a child's environment. This was partly driven by the relationships between behaviour and primarily social factors in the environment such as family relationships, caregiver wellbeing and skills and skills available to the family via acquaintances. Secondly, there was no association between cognition and behaviour that could not be explained by differences in a child's environment. It is possible that a better environment caused both increased cognitive skills and better behaviour, and as a result they only appeared to be associated, a key issue in correlational studies.

This is interesting given the large body of literature that identified an association between cognition and behaviour (Flouri et al., 2015; Harpur et al., 2015; Helton et al., 2018; Hinshaw, 1992; Hughes and Ensor, 2011; Masten et al., 1999; Menting et al., 2011; Willoughby et al., 2011). In particular, the BRIEF subscales are designed to be a behavioural index of executive functions, such as working memory. However, the assumption that the BRIEF scale measures behaviours that are closely related to cognitive processes assessed by performance-based tasks of executive function has recently been called into question. For example, in their review of the literature measuring the correlations between different executive function tasks and BRIEF ratings, Toplak et al. found that of the 182 correlations reported, only 19% had significant correlations with a mean correlation of 0.15 (Toplak et al., 2013). They also suggest that this is an overestimate due to type 1 errors and may be additionally inflated due to publication bias. They hypothesised that ratings of behaviour and performance

testing of executive functions measure different aspects of cognitive and behavioural functioning. If this is the case, we would expect the environment domains to explain different variance in each dataset.

It is also possible that these results could be due to the cognitive tasks used. In particular, we did not use any tasks that might tap 'hot' executive functions, those described as affective or related to emotions. If we consider the correlation between the cognitive subtests and the different subscales in the BRIEF and SDQ (found in Appendix E, Figure 90), we found that around half of the behaviour subscales were moderately correlated with the cognitive subtests administered. The behaviour subscales that are not correlated are primarily the scales relating to 'hot' functions such as BRIEF Emotional Control, SDQ Emotion Problems or Peer Problems. Furthermore, the environmental domains identified to covary with behaviour and not the cognitive tasks measured are also primarily related to social measures. Note that the executive functions assessed in the review by Toplak et al. were all 'cold' skills and there is evidence in the literature that behaviour may be more strongly related to 'hot' skills. For example, whilst Willoughby et al. found that academic abilities were closely associated with 'cold' and not 'hot' self-regulation skills, the reverse was largely true for disruptive behaviours (Willoughby et al., 2011). In another example, it was found that impulsivity and not sustained attention, typically considered 'hot' and 'cold' skills respectively, was found to mediate the relationship between the quality of the home environment and social competence, and externalizing behaviours (NICHD Early Child Care Research Network, 2003). Therefore, whilst we found that the cognitive skills assessed do not relate to behaviour when differences in the environment are accounted for, it is possible that other tasks such as affective executive function tasks may relate more strongly to behaviour outcomes and we cannot rule out the possibility that other cognitive skills may provide a potential mechanism for the environment to impact a child's behaviour.

Finally, it is also possible that the environment may instead mediate the relationship between cognition and behaviour, reversing the path diagram. A key issue of cross-sectional mediation analyses such as this is that the direction of the arrows must be estimated from theory, which are assumed to be correct rather than confirmed by the analysis. Whilst we hypothesized that the environment may impact the development of cognitive functions which in turn may affect the outcomes, it is possible that cognitive skills may influence some of the environment domains. For example, if a child has higher cognitive abilities, they might be more likely to read at home more or require

less discipline. When we reverse the model the indirect path between cognition and behaviour via the environment is significant (First component ab=0.21±0.05, 95% CI [0.17,0.36]). However, there are several environmental variables that cannot be hypothetically caused by child cognition (e.g. child health, parent income, education or occupation). Therefore, we would argue that conclusions cannot be drawn from this significant indirect effect.

In summary, we have built on the body of literature examining pairwise relationships between environmental domains and cognitive skills by using PLS methods to explore how multiple environmental and cognitive factors work together in concert. We have identified the set of environmental domains that are most strongly associated with a child's cognitive profile. Many of these environmental domains mediated the apparent relationship between standard SES measures and cognition. Furthermore, we have demonstrated that a child's cognition mediates the relationship between their environment and academic outcomes but not their behaviour outcomes. Instead, we found that the environment not only explains additional variance in the behavioural outcomes over and above variance explained by cognition, but also fully explains the association between the cognitive skills and behavioural outcomes.

5 ENVIRONMENT AND THE STRUCTURAL CONNECTOME

5.1 Introduction

The structural connectivity of the human brain changes profoundly from birth to early adulthood. Developments in microscopic structure, such as synaptic connectivity and axonal myelination are reflected in macroscopic changes in the volume and architecture of white matter, the axons that connect regions of grey matter (Barnea-Goraly et al., 2005; Muftuler et al., 2012; Vértes and Bullmore, 2015). Concurrently, the pattern of connectivity shifts, with the emergence of distributed brain networks, gradually replacing connectivity patterns dominated by local connections (Cao et al., 2014; Fair et al., 2009; Richmond et al., 2016). These changes in structural connectivity are thought to underpin the development of a broad range of abilities such as a child's behavioural control and academic and cognitive abilities which typically rely on the integration of a number of regions across the brain (Pugh et al., 2001; Tau and Peterson, 2010).

Structural connectivity is often probed using the diffusion properties of water molecules, measured by diffusion tensor imaging (DTI). In structured spaces, such as within white matter, water is constrained to move along the axon as diffusion across the axon wall is restricted. These diffusion properties reflect the structural integrity of white matter. As this structure influences how effectively information can be transmitted between brain regions, individual differences in white matter integrity are of particular interest for child development (Fields, 2008; Johansen-Berg, 2010; Scholz et al., 2009). In particular, fractional anisotropy (FA), the directionality of water diffusion, is commonly used in neuroimaging studies as a summary indicator of white matter integrity that reflects fibre density, axonal diameter and myelination. FA increases in widespread regions across the brain during childhood and adolescence (Giorgio et al., 2008, 2010; Lebel et al., 2017; Snook et al., 2005). For example, FA has been found to be a more sensitive measure of brain development than grey matter cortical thickness in a large sample of 6-16 year olds (Bathelt et al., 2018b). Widespread differences in FA are associated with a range of developmental markers such as reading (de Moura et al.,

2016; Deutsch et al., 2005; Klingberg et al., 2000; Qiu et al., 2008; Vandermosten et al., 2012), maths (Bathelt et al., 2018a; Matejko and Ansari, 2015; Tsang et al., 2009; Van Beek et al., 2014), behavioural difficulties (Loe et al., 2013; Muetzel et al., 2018; Waller et al., 2017) and cognitive skills (Bathelt et al., 2018b; Mabbott et al., 2006; Nagy et al., 2004; Roberts et al., 2013; Takeuchi et al., 2011; Tuch et al., 2005; Ursache et al., 2016; Vestergaard et al., 2011). Furthermore, there is evidence that white matter structure changes due to training in particular skills. For example, FA increases were found after training in a complex visuospatial coordination task (juggling) (Scholz et al., 2009), working memory (Takeuchi et al., 2010) and reading training (Keller and Just, 2009) and after even short periods of meditation (Tang et al., 2010).

Research investigating the relationship between the environment and brain structure has primarily been focused on grey matter differences, and little is known about the relationship of the environment to FA. There is some evidence that heritability of FA appears to decrease with age, in contrast to white matter volume which tends to have stable high heritability across development (Chiang et al., 2011; Douet et al., 2014; Richmond et al., 2016). This suggests that FA might be particularly susceptible to differences in the environment that children experience. Adult SES measures have been found to directly relate to FA in regions across the brain (Gianaros et al., 2013; Noble et al., 2013). In a study by Gullick et al. of 32 7-11 year old children using tract based spatial statistics (TBSS), FA was found to be positively associated with parental SES in clusters across the brain: the left corticospinal tract, right anterior inferior fronto-occipital fasciculus, left superior longitudinal fasciculus and left temporal inferior longitudinal fasciculus (Gullick et al., 2016). No negative associations were found. Similar results were found in an another TBSS analysis by Dufford and Kim of 27 8-10 year old children, in which higher FA in the uncinate fasciculus, cingulum bundle, inferior longitudinal fasciculus, superior longitudinal fasciculus, and corticospinal tracts was found to significantly related to parent income (Dufford and Kim, 2017). In a much larger study of 1082 3-21 year olds, Ursache et al. found that higher parental income and education was associated with higher FA in ROIs in the cingulum bundle and the superior corticostriate tract in the frontal and parietal cortex (Ursache et al., 2016). In contrast, Jednorog et al. found in a sample of 23 10 year old children that parental SES (education and occupation) was related to grey matter volume, surface area and gyrification but found no relationship to FA using TBSS (Jednorog et al., 2012). Overall, these studies indicate that higher SES is generally related to higher FA values

in regions across the brain. However, given the limited research, generally low sample sizes and inconsistent approaches and results, there is a great need for further research.

Even fewer studies consider the environment beyond the standard SES measures. Dufford et al. found that a cumulative risk factor summarizing house crowding, noise, quality and family turmoil, violence, and child separation were associated with FA in regions across the brain (Dufford and Kim, 2017). Gianaros et al. found in a sample of adults that waist circumference and smoking mediated the relationship between SES measures (including neighbourhood SES) and FA (Gianaros et al., 2013). Finally, 7 children subjected to early environmental deprivation in Romanian orphanages had lower FA values in the left uncinate fasciculus in comparison to healthy controls (Eluvathingal et al., 2006). The previous chapters have identified multiple aspects of a child's environment that are related to childhood development. It is likely that many of these are also related to brain structure and there is a great need to identify the specific aspects of the environment that are most influential in explaining differences in FA.

In addition, little work has been done to investigate whether differences in brain structure underlie the relationship between environmental factors and developmental outcomes. The study by Gullick et al. found that SES moderated the relationship between FA and reading skill in children. FA was found to mediate the relationship between educational attainment and cognitive control in 17-23 year olds (Noble et al., 2013). These effects were reflected in other structural differences. For example, in a large longitudinal study of children and adolescents, Hair et al. found that development of grey matter volumes in the frontal and temporal lobes and the hippocampus accounted for 15-20% of the income related academic attainment gap. Finally, Brito et al. found that cortical surface area mediates the relationship between income and cognitive skills (inhibitory control and working memory) (Brito et al., 2017). Given the importance of communication between brain regions for academic and behaviour outcomes, the impact of the environment on white matter integrity holds potential as a mechanism for the environment to exert an effect on these developmental outcomes.

Furthermore, the majority of developmental studies have focussed on local differences in FA using univariate voxel-based analyses such as comparisons between ROIs or using tract based spatial statistics (TBSS). These studies have been critical in demonstrating the associations between individual differences in white mater integrity and both environmental factors and child development outcomes. However, child outcomes such as academic ability and behaviour require the co-ordinated integration of

a number of widespread brain regions. Furthermore, in general, differences in both child outcomes and SES are related to differences in FA in white matter distributed across the brain. Voxel-wise methods that rely on overlapping associations between FA and other variables in specific voxels across children are likely to underestimate the contribution of whole-brain network structure that integrates the distributed brain systems (Bathelt et al., 2018a). This is because these methods are primarily designed to identify highly focal and consistent differences across individuals. By contrast, differences that emerge over development are unlikely to be focal, because these differences will cascade over time, having consequences for processing in other parts of the system and potentially being compensated for elsewhere. Voxel-wise methods are unable to detect these effects. For example, Bathelt et al. contrasted the global organization of a child's FA connectome with the FA skeleton found using TBSS. They found that TBSS was far less sensitive to a child's academic development than the connectome methods (Bathelt et al., 2018a). Rather than considering each voxel or group of voxels in turn, we will consider how whole-brain connectivity is associated with the environment and child development by calculating the FA along the length of the white matter connections (tracts) between each pair of regions in the brain. This will allow us to further investigate whether associations between the environment and a child's brain development are localised to specific brain regions or whether there is a more general relationship distributed across the brain architecture.

We can build up a map of these tracts by calculating the diffusivity of water in multiple directions and use this to estimate the orientation of axon bundles (tractography) (Hagmann et al., 2007), allowing us to map the pathways of the human brain at unprecedented resolution. Using this, a graph model of the brain can be created by parcelling the grey matter into regions (nodes), often using T1-weighted MRI images, and considering the tracts (edges) that intersect with each pair of nodes (Hagmann et al., 2007; Zalesky et al., 2010). The average FA along each edge can then be calculated to provide a measure of the white matter integrity joining each pair of grey matter regions in the brain. Paired with multivariate methods such as PLS, we can use this to investigate the association between the structural integrity of network connections, the environment and child development across the whole brain (McIntosh and Lobaugh, 2004). For example, a similar method has been used to demonstrate the relationship between a large number of environmental variables to the functional connectome in adults (Smith et al., 2015). In another adult study, the connectome based on similarities

in cortical thickness between brain regions was found to differ based on neighbourhood SES (Krishnadas et al., 2013). However, to our knowledge, no other study has considered the relationship between the environment and any connectome measures in children, the relationship between the environment and the FA connectome at any age nor whether this relationship mediates environment-related differences in child developmental outcomes.

In order to address the issues outlined, we used PLS methods to explore the associations between environmental variables and FA in a child's structural connectome to answer three questions. First, what are the key aspects of a child's environment most strongly associated with their structural connectome? Second, does the wider environment mediate the relationship between the standard SES measures and their structural connectome? Finally, does FA act as a potential mediating mechanism underlying the relationship between environment and academic and behaviour outcomes?

5.2 Methods

5.2.1 Participants and structural MRI

Diffusion Tensor Images (DTI) and T1-weighted images were acquired for 86 children (10 children opted out of the MRI and 1 child did not complete the DTI scan) at the MRC Cognition and Brain Sciences Unit, Cambridge, U.K. A total of 17 scans were excluded: 10 due to large movement (maximum displacement above 3mm in the DTI sequence, as calculated by FSL Eddy), 4 due to scanner errors. A further three participants were excluded due to having a large percentage of missing questionnaire data. The resulting sample included 69 children (32 males, mean age 9:11 and age range 6:11-12:9). Scans were conducted using a 3T Tim Trio system (Siemens Healthcare, Erlangen, Germany) and a 32-channel quadrature head coil. A Magnetisation Prepared Rapid Acquisition Gradient Echo (MP RAGE) sequence with 1mm isometric image resolution, 2.98ms echo time and 2250ms repetition time was used to acquire whole brain T1-weighted volume scans. An Echo-planar diffusion-weighted sequence with an isotropic set of 60 non-collinear directions, using a weighting factor of b=1000s*mm-2, interleaved with 4 T2-weighted (b = 0) volumes, 60 contiguous axial slices, isometric image resolution of 2mm, 90ms echo time and 8400ms repetition time was used to acquire the diffusion weighted images. Children watched a film of their choice whilst undergoing the 20-minute MRI scan.

5.2.2 The structural connectome

The white-matter DTI connectome was constructed by estimating the most probable white matter connections for each participant and then constructing connectivity matrices from the average Fractional Anisotropy (FA) between each pair of brain regions.

An overview of the pre-processing steps is given in Figure 37. Raw MRI scans were first converted from DICOM to compressed NIfTI-1 format using the dcm2niitool (http://www.mccauslandcenter.sc.edu/mricro/mricron/dcm2nii.html). The b_0-weighted volume of the DTI images was used to create a brain mask and all volumes were corrected for movement and eddy currents using FSL's eddy tool. Following this, nonlocal means denoising (Coupe et al., 2008) was used to improve the signal-to-noise ratio using the Diffusion Imaging in Python (DiPy) v0.11 package (Garyfallidis et al., 2014). Fractional Anisotropy maps were derived for each participant by fitting the diffusion tensor model using dtifit from the FMRIB Software Library (FSL) v.5.0.6 (Behrens et al., 2003). DiPy was then used to fit a spherical constrained deconvolution (CSD) (Tournier et al., 2008) model to the 60-gradient direction DTI images using a maximum harmonic order of 8. Subsequently, probabilistic whole-brain tractography was applied to this CSD model using 8 seeds in any voxel that had a General FA value greater than 2, step size equal to 0.5 and with no more than 2 crossing fibres allowed per voxel.
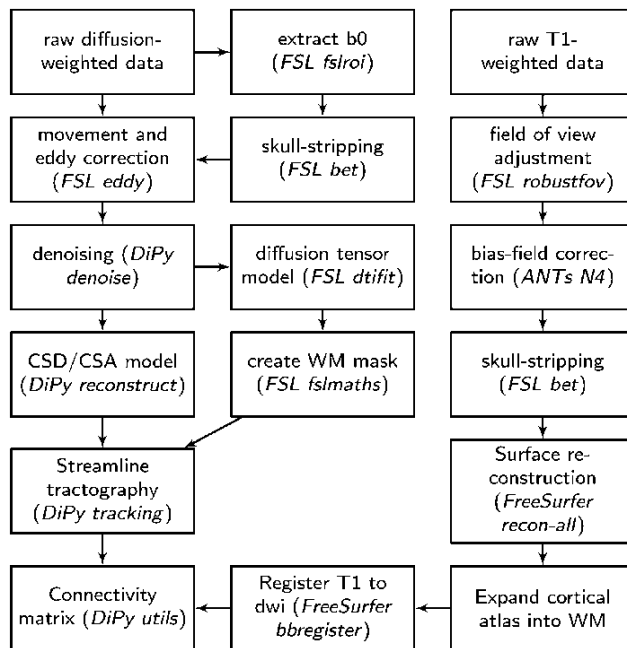
Figure 37 Overview of the pre-processing steps used to derive the DTI connectome. Reprinted from 'Whole-brain white matter organization, intelligence, and educational attainment' by Bathelt et al., 2018, bioRxiv preprint, doi: http://dx.doi.org/10.1101/297713 with permission.

The T1-weighted images were used to define the brain regions. The images were first pre-processed by adjusting the field of view using FSL's robustfov function and denoised using nonlocal means denoising in DiPy. A robust brain mask was extracted using the brain extraction algorithm of the Advanced Normalization Tools (ANTs) v1.9 package (Avants et al., 2009). The images were then submitted to the recon-all pipeline in FreeSurfer v5.3 (http://surfer.nmr.mgh.harvard.edu). 85 ROI's were extracted using the Desikan-Killiany parcellation of the MNI template (Desikan et al., 2006). 34 regions were extracted for each hemisphere and 17 subcortical regions (brain stem, and bilateral cerebellum, thalamus, caudate, puta-men, pallidum, hippocampus, amygdala, nucleus accumbens). The aparc2aseg tool in FreeSurfer was used to convert the surface parcellation of the cortex to a volume parcellation. Following this, the cortical parcellation was expanded by 2 mm into the subcortical white matter to ensure that the tracts would intersect the ROIs using in-house software. Finally, the parcellation was transformed into DTI space using a transformation based on the T1-weighted volume and the b0-weighted image of the diffusion sequence calculated using FreeSurfer's bbregister.

FA-weighted connection matrices were created for each participant by considering the FA values across the streamlines from the CSD model that connected each pair of ROIs. The corresponding element in the connection matrix for each pair of ROIs was either set to zero if no streamlines intersected both ROIs, or to the average FA of the streamlines intersecting both ROIs. In order to remove spurious connections, a common problem in connectome studies, the matrices were first thresholded using consensus thresholding (de Reus and van den Heuvel, 2013). Consensus thresholding retains connections that are found in a minimum percentage of participants determined by a given threshold. A consensus threshold of 0.6 was used, based on recommendations given by (de Reus and van den Heuvel, 2013), ensuring that only connections that are found in over 60% of the sample are retained. This resulted in 265 non-zero FA weighted connections between brain ROIs (see Figure 38). Finally, the matrices for each subject were reshaped into a single vector and combined across subjects, resulting in a 265 X 69 dataset in which each column represents the FA between each pair of regions for each of the 69 participants.
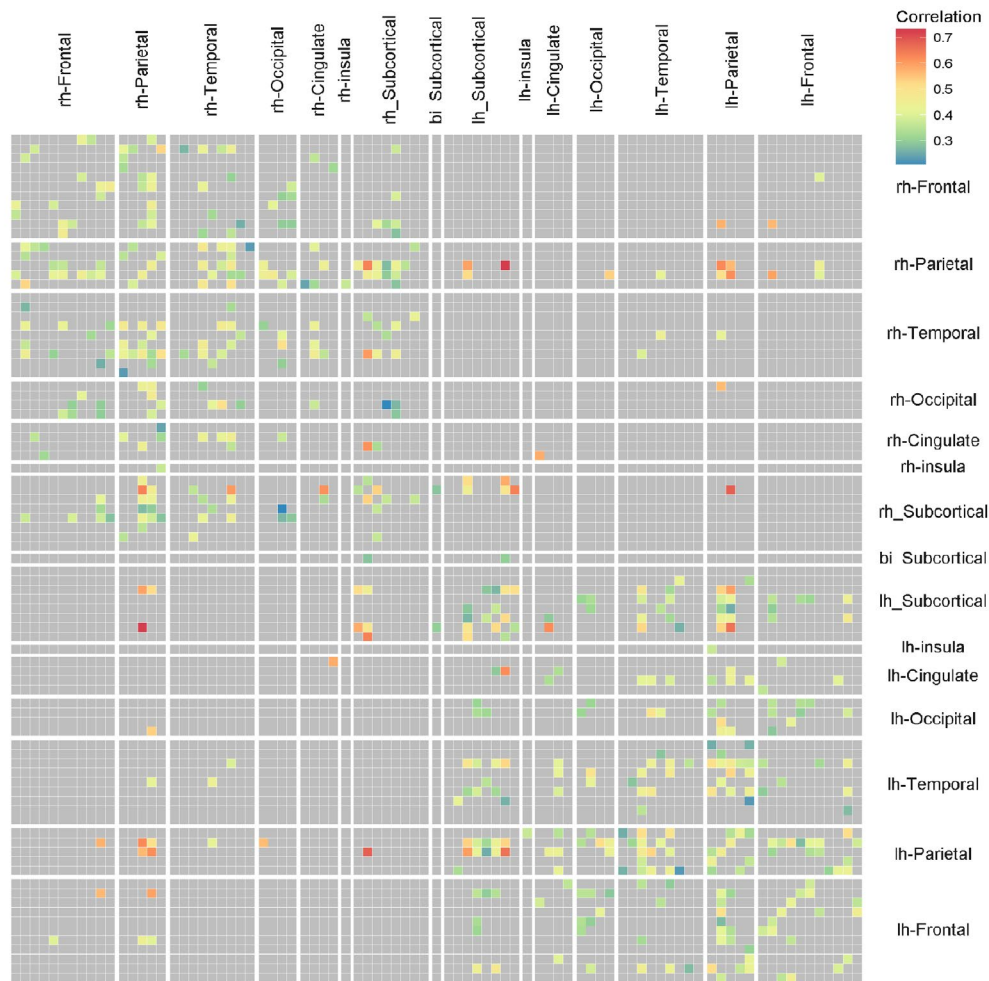
Figure 38 Group-average connectome matrix indicating the ROI-by-ROI FA weighted connections, after 60% consensus thresholding. The ROIs are grouped by lobe and hemisphere (lh= left hemisphere, rh= right hemisphere, bh= bilateral hemispheres).

5.2.3 PLS between the environment and the structural connectome

PLS is ideally suited to summarising the complex relationships between connectomes and other datasets due to the large number of often multicollinear variables in connectomics. PLS was used to identify the profile of environmental factors that most strongly covaried with the FA connectome. Each variable was standardised and the PLS analysis was conducted as in section 3.2.2. The environmental variables and connectome data were used as dataset X and Y respectively. Rubin's rules were used to pool the PLS results across imputation and the significance and reliability of the PLS model were assessed using permutation (N permutations = 1000) and bootstrap (N bootstrap samples = 500) analyses.

5.2.4 PLS mediation

We applied 3-block Canonical PLS to the standardised datasets as detailed in section 3.2.3 to investigate mediation effects. First, the mediating effect of the wider environment on the SES-connectome relationship was investigated using the three standard SES measures, equivalised income, caregiver education and occupation as dataset X, the other environmental domains as dataset M and the connectome data was used as dataset Y. Note that, as in section 3.2.3, primary caregiver's hours work per week was excluded from the wider environment dataset.

Secondly, to investigate whether a child's brain structure mediated the relationship between a child's environment and their academic and behaviour outcomes, the full set of environmental domains were used as dataset X, the connectome was used as dataset M and either the two academic variables or two behaviour variables were used as dataset Y.

5.3 Results

5.3.1 PLS: What is the environmental profile that most strongly relates to a child's structural connectome?

Two-block PLS was conducted using the environmental and structural connectome datasets. The first pair of latent variables significantly explain the covariance between the environmental domains and the structural connectome (p = 0.032, explained variance: 20% environment and 8% structural connectome, average correlation between environment and structural connectome latent variable = 0.69). In general, high FA was related to positive environmental domain scores. Several environment domains were found to reliably covary with the structural connectome (Figure 39). Notably, all three key standard SES measures, occupation, income and education are found to be positively related to high FA values. In addition, a number of the other measures selected were strongly related to SES including the number of hours the primary caregiver works, subjective SES and neighbourhood SES. In addition, child health, attitude to child education, number of caregivers and childcare were each positively related to high FA values. The brain connections that covaried reliably with the environment latent variable are widespread across the brain (Figure 40). In general, these connections were longer-range connections between different lobes rather than within a lobe. The second component was not significant.
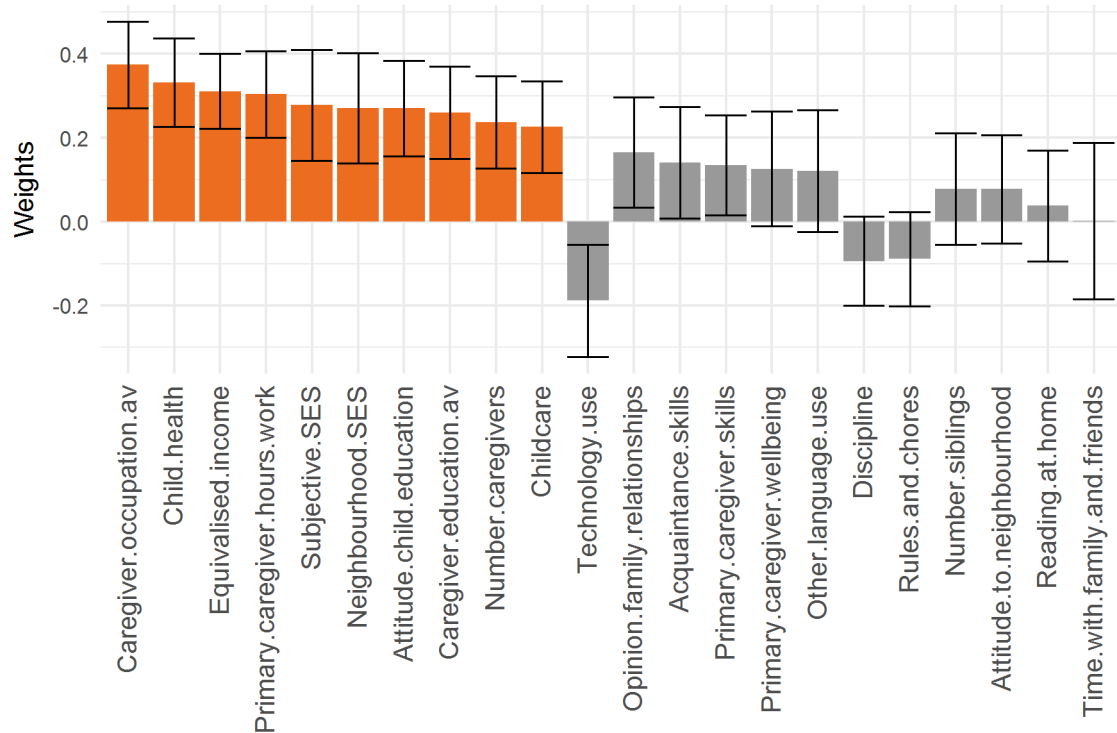
Figure 39 The environment and the structural connectome: outer weights for the first latent variables, averaged across imputations . Error bars are the standard errors obtained from the bootstrap samples, pooled across imputations using Rubin's rules. Reliable loadings (according to the critical t-value) are coloured orange.
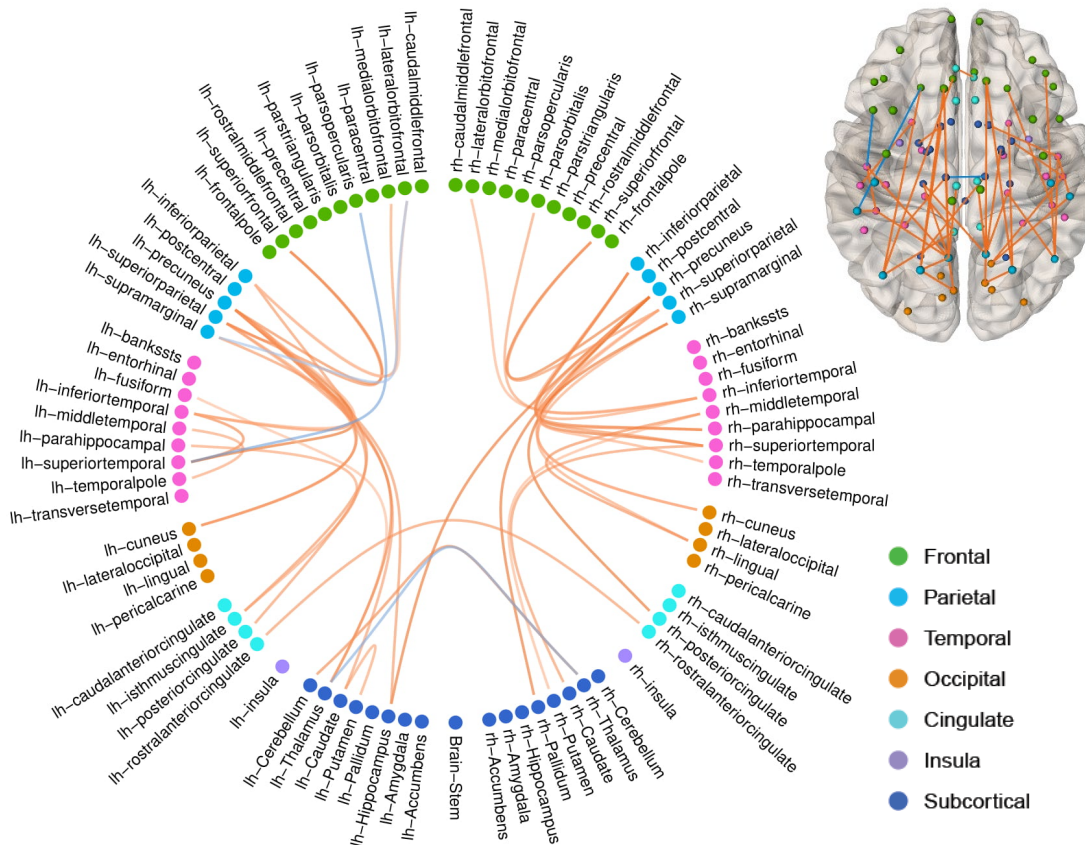
Figure 40 The environment and the structural connectome: a topographic and circle plot of the edges that were reliably related to the environmental domains based on the bootstrap of the first PLS component, coloured by the sign of the loading onto the first component (orange, positive; blue, negative). Edges are grouped by brain hemisphere (L/R) and lobe (given in the key).

5.3.2 PLS Mediation

5.3.2.1 Does the wider environment mediate the relationship between standard SES measures and the structural connectome?

The wider environment significantly mediated the relationship between SES and the structural connectome for the first latent variable (indirect path coefficient $ab=0.31\pm0.09$, 95% CI [0.25,0.60]). The path model for the first component can be seen in Figure 41 and the variable weights can be seen in figure Figure 42. Note that the direct path between SES measures and the structural connectome remains significant even when the other environmental domains are accounted for, indicating that the SES measures explained unique variance in the structural connectome that was not related to the wider environment domains measured. The second component was significant and the third was marginally significant.
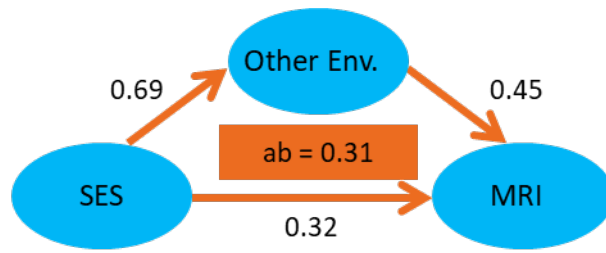
Figure 41 PLS mediation analysis between SES, the wider environment and the structural connectome: the path model
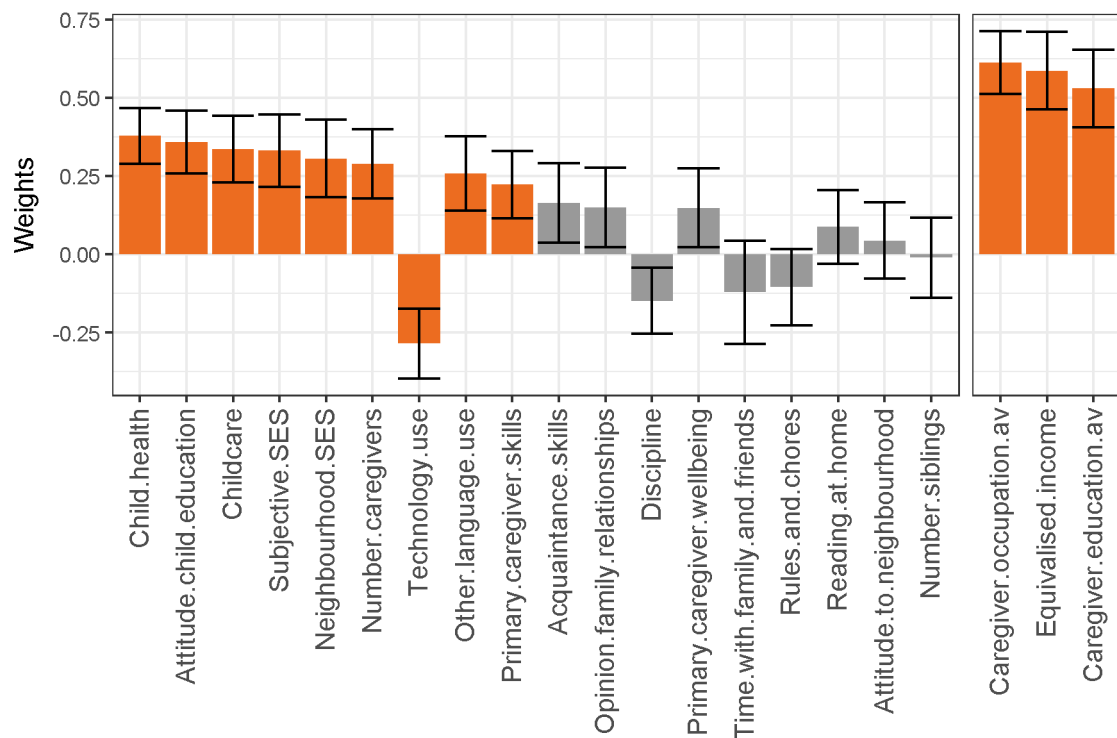
Figure 42 PLS mediation analysis between SES, the wider environment and the structural connectome: outer weights for the first latent variables, averaged across imputations
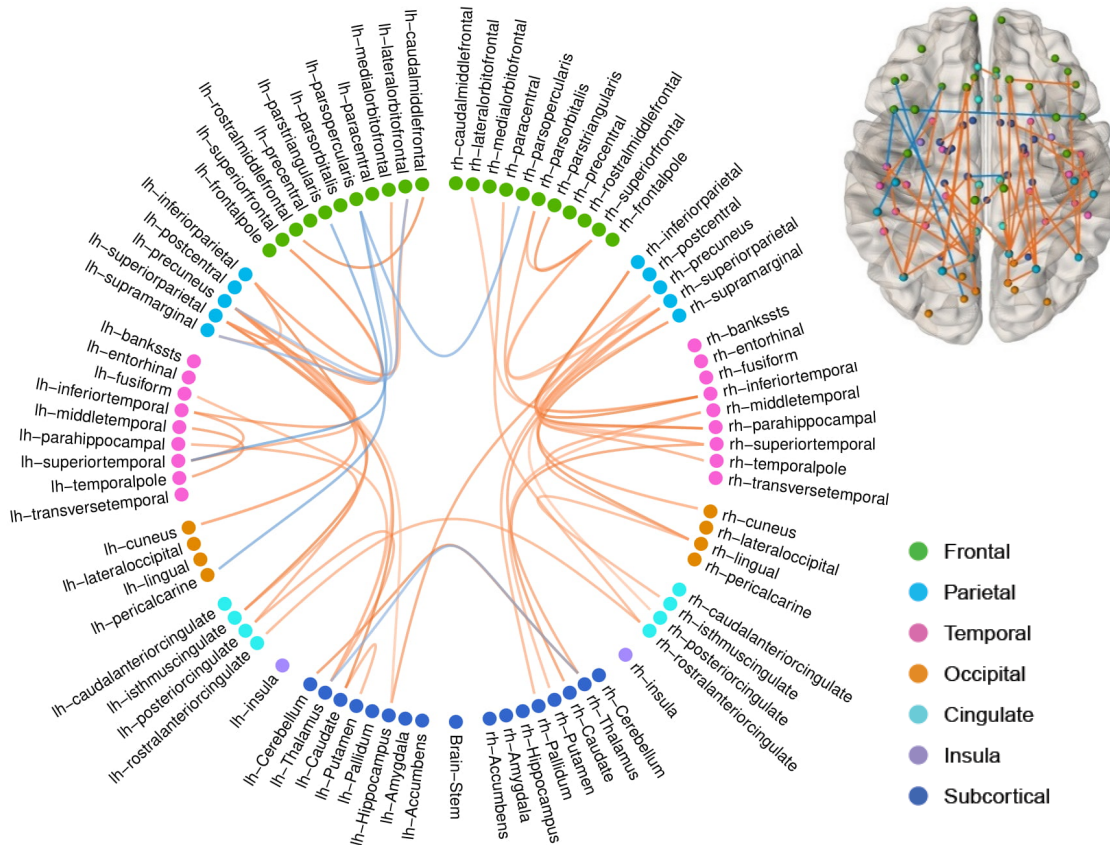
Figure 43 PLS mediation analysis between SES, the wider environment and the structural connectome: a topographic and circle plot of the edges that loaded reliably onto the latent variable.

Each of the standard SES measures loaded relatively equally strongly onto the SES latent variable. The wider domains that loaded onto the mediator latent variable included those that were identified as reliable covariates of the structural connectome in 5.3.1. In addition, inverse technology use, other language use and primary caregivers are also found to load onto the mediation latent variable. The pattern of connections in the structural connectome had a similar profile to the profile found for the environment-structural connectome relationship, although there were also some additional connections that were found to reliably load onto this profile.

5.3.2.2 Does the structural connectome mediate the relationship between a child's environment and academic ability?

The structural connectome fully mediated the relationship between the environment and academic outcomes in the first component (indirect path coefficient ab=0.27±0.16, 95% CI [0.16,0.80]). The path model can be seen in Figure 44 and the variable weights can be seen in Figure 45. Reading and maths loaded with equal strength onto the academic

latent variable. Caregiver occupation, child health, attitude to education, caregiver education, primary caregiver work hours, number of caregivers, equivalised income and childcare each loaded reliably onto the environment latent variable. Finally, the outer weights for the individual connections between brain regions were almost entirely positive. Note that the mediation effect for the second component was also significant but the relationship between the environment and academic second components was not significant, making this difficult to interpret.
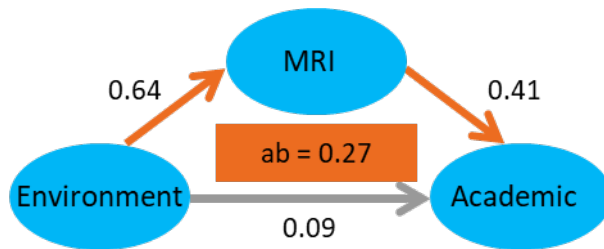
Figure 44 PLS mediation analysis between the environment, the structural connectome and academic ability: the path model
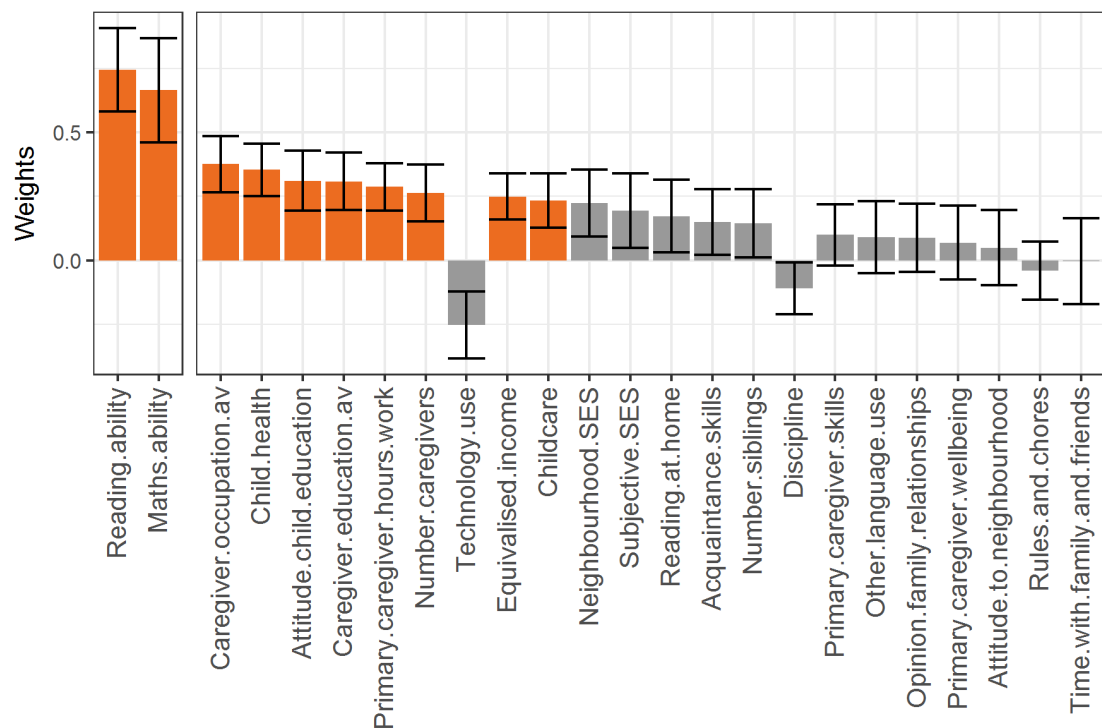
Figure 45 PLS mediation analysis between the environment, the structural connectome and academic ability: outer weights for the first latent variables, averaged across imputations
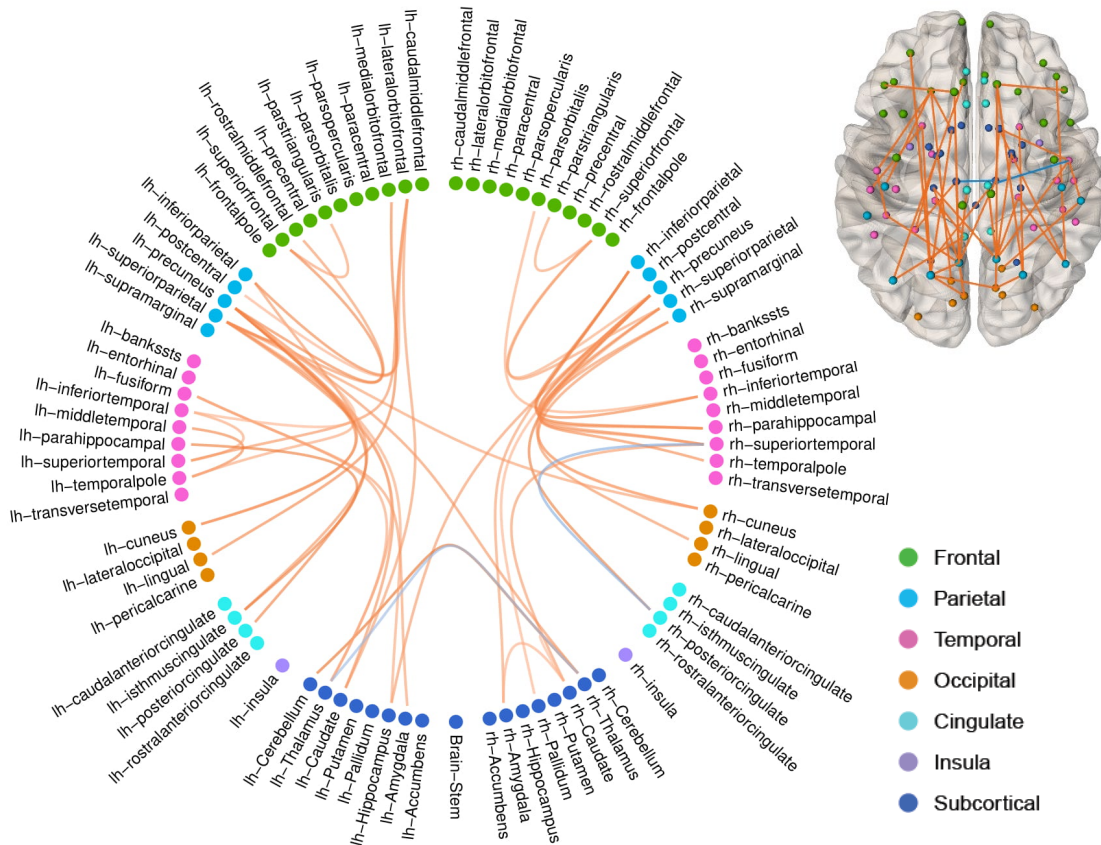
Figure 46 PLS mediation analysis between the environment, the structural connectome and academic ability: a topographic and circle plot of the edges that were reliably related to the latent variable

5.3.2.3 Does the structural connectome mediate the relationship between a child's environment and behaviour?

The structural connectome did not mediate the relationship between the environment and behaviour outcomes in the first component (indirect path coefficient ab=0.16±0.22, 95% CI [-0.07,0.78]). The path diagram is given in Figure 47 and the weights that reliably load onto each latent variable can be seen in figure Figure 48.
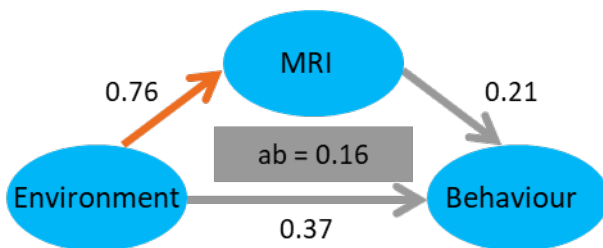
Figure 47 PLS mediation analysis between the environment, the structural connectome and behaviour: the path model
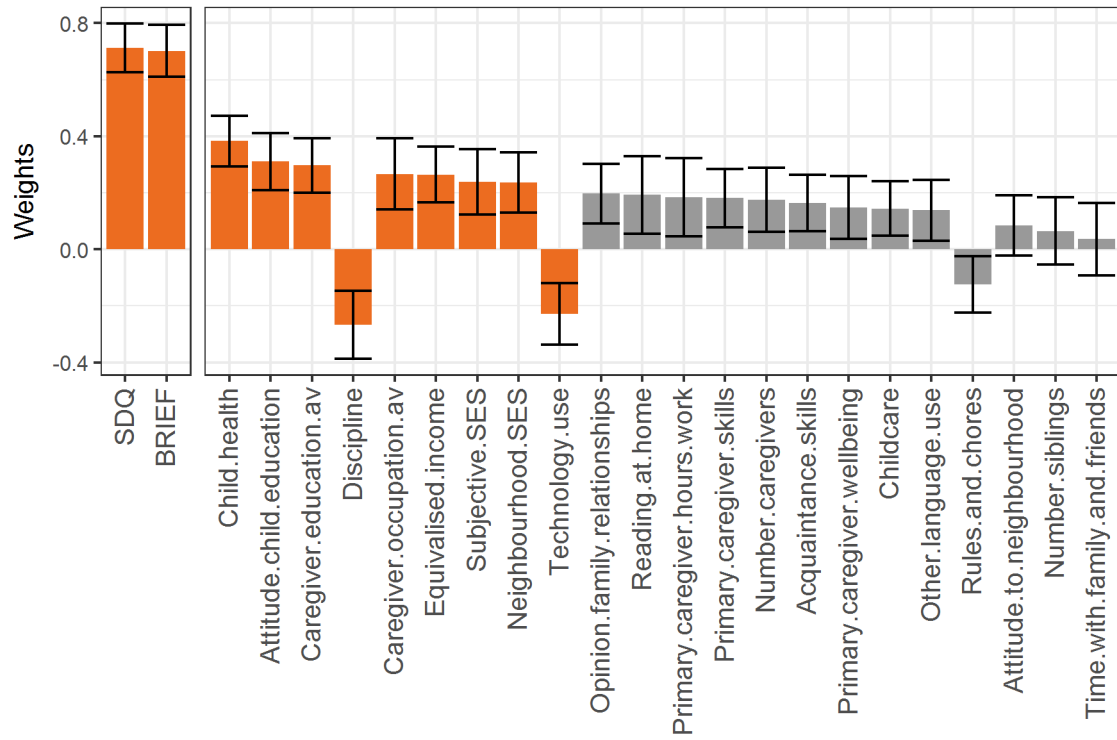
Figure 48 PLS mediation analysis between the environment, the structural connectome and behaviour: outer weights for the first latent variables, averaged across imputations
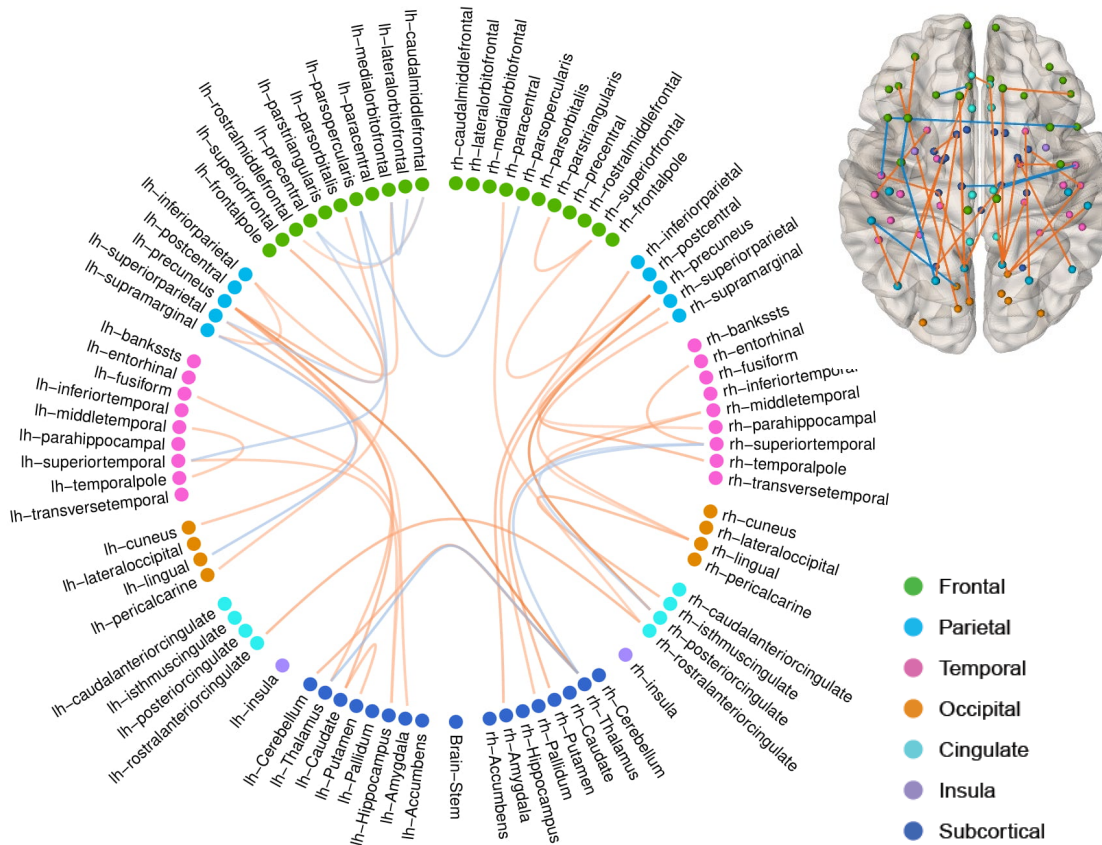
Figure 49 PLS mediation analysis between the environment, the structural connectome and behaviour: a topographic and circle plot of the edges that were reliably related to the latent variable).

5.4 Discussion

Previous studies using voxel-wise comparisons have demonstrated localised relationships between SES and white matter integrity globally and at ROI's across the brain (Gianaros et al., 2013; Noble et al., 2013). These have had big theoretical implications within this field, because these previous results imply that SES impacts upon highly specific brain mechanisms and outcomes. We have built on these findings by investigating the structural integrity of white matter across the whole-brain connectome and its association to multiple aspects of a child's environment and developmental outcomes. To our knowledge, this is the first study to investigate the relationship between the environment and the FA connectome (or indeed any type of connectome in children).

5.4.1 The environment and the structural connectome

The results demonstrate that, in general, FA is positively related to multiple positive aspects of a child's environment. The brain connections that reliably covary with the environment latent variable are widely distributed across the brain. The parietal and temporal lobes are particularly well connected with connections to the frontal, occipital, cingulate cortices and subcortical areas. There are also temporo-parietal connections and connections between the left and right cingulum and left and right subcortical regions. Interestingly, the reliable connections across the left hemisphere are highly similar to those selected in the right hemisphere.

These results align with the voxel-wise findings that SES is associated to FA in various white matter tracts connecting different lobes of the brain (Gullick et al., 2016; Noble et al., 2013; Ursache et al., 2016). However, whole-brain connectomics allows us to demonstrate that, rather than being limited to effects in focal ROIs, the association is remarkably widespread, suggesting a more global relationship between the environment and FA. This is likely to be due to the multifaceted nature of the environment (Raizada and Kishiyama, 2010), and it may be that different aspects of the environment contribute to explaining FA variability in different tracts across the brain. For example, differences in child cognitive stimulation (such as caregiver education or attitude to education) will result in more or less stimulation of particular connections which, in turn, can alter the myelination around the neurons (Demerens et al., 1996). In addition, the developing brain is a highly dynamic and interactive system. Differences in one region or connection between regions are likely to generate further differences in other areas, resulting in cascading effects across the brain (Johnson, 2011). It is also possible that a generally poorer environment may contribute to a systematic reduction in FA, such as an association with fibre density, diameter or myelination at a global scale. For example, it has been found that exposure to excessive levels of stress hormones can suppress glial cell division which is critical for myelination (Lauder, 1983). In another example, increased BMI was associated with decreased global FA and in regions across the brain and this was partially explained by differences in vascular risk factors, blood markers of inflammation and vascular health (Bettcher et al., 2013; Verstynen et al., 2012). These inflammatory pathways have been found to mediate the relationship between SES and FA (Gianaros et al., 2013). In turn, widespread differences in FA are likely to translate into multiple behavioural differences, providing a possible

explanation as to why such a wide variety of outcomes are strongly associated with a child's environment.

The environment domains that most strongly covaried with the structural connectome were: caregiver occupation, child health, equivalised income, primary caregiver work hours, subjective SES, neighbourhood SES, attitude to child education, caregiver education, number of caregivers, childcare. It is particularly interesting that the standard SES measures and domains thought to be closely related to these measures such as subjective and neighbourhood SES and work hours dominate the environment latent variable. This is in contrast to the previous results for academic ability, behaviour and cognitive skills in which, with the exception of caregiver education and occupation, none of these environmental domains were strongly associated with child development outcomes. Notably, equivalised income, which does not correlate with any of the academic, behaviour or cognitive scores was one of the domains most closely related to the FA connectome. There are a few possible explanations for this. For example, it might be the case that SES differences in FA precede observable differences in child outcomes at this age. For example, low income may result in poor development of white matter, but this may only be reflected in outcome scores at a later age. That is, different environmental factors operate over different time-scales, and our neural and outcome measures are differentially sensitive to these time-scales. Alternatively, there may be processes that protect children from the influence of low FA on developmental outcomes. For example, the wider environmental domains may act to buffer the negative relation between SES-related low FA and child outcomes (Ursache et al., 2016), and this is why they are most closely aligned with the outcome measures.

Furthermore, the wider environment was found to significantly mediate the relationship between SES and the structural connectome. Each of the environment domains highlighted as covariates of FA, child health, attitude to education, childcare, subjective SES, neighbourhood SES and number of caregivers, loaded reliably onto the mediating latent variable, suggesting that these hold potential as mechanisms that transfer the effect of SES to FA. However, the SES measures also reliably explain additional unique variance in the FA connectome that is unrelated to the wider environment domains. Although correlational, this suggests that there may be other mechanisms by which SES might affect FA across the brain that are not captured in the domains measured.

5.4.2 The structural connectome mediated the relationship between the environment and academic outcomes

Studies have reported associations between FA and academic skills and between the environment and FA, suggesting that FA might partially explain links between the environment and academic outcomes. However, given the fact that several of the environmental domains found to covary with FA did not covary with the academic outcomes, it is perhaps surprising that FA fully mediated the relationship between the environment and academic ability. This is likely to be driven by the environmental domains that did covary with both: caregiver education and occupation, child health, attitude to child education and childcare. One possible interpretation of these results is that a better childhood environment might facilitate positive structural development in a child's brain necessary for the development of academic skills.

Note that the connections that were most strongly related to this mediation effect were widespread across the brain. However, the outer weight for each connection is the sum covariance between the tract FA and both the environment and academic datasets. Therefore, it is difficult to know whether all of these connections are related to academic ability or whether some weights are driven primarily by high covariance with the SES measures such as income.

Currently, there are very few studies to draw on illustrating potential causality behind these correlations. Only one longitudinal study is listed in the recent review of white matter and reading literature (Vandermosten et al., 2012) and no longitudinal studies are described in the review of maths ability (Matejko and Ansari, 2015). Hoeft et al. found that FA in the right superior longitudinal fasciculus predicted improvement in reading over 2.5 years in children diagnosed with Dyslexia (Hoeft et al., 2011). Note that this association was not found in children without dyslexia. On a shorter timescale, Floel et al. found that FA in tracts passing through Broca's area predicted success in an artificial grammar learning task in adults (Flöel et al., 2009). Finally, FA increases have been found after reading training. Keller and Just found that poor readers had significantly lower FA in a region in the left frontal white matter in comparison to good readers (Keller and Just, 2009). Following 100 hours of intensive reading instruction over a 6-month period, poor readers demonstrated an increase in FA in the same region in comparison to a control group of poor readers who received normal classroom reading training. Furthermore, the increase in FA correlated with improvements in phonological decoding ability. These studies provide initial evidence towards a causal link between

reading and FA, but further research is needed to demonstrate a causal link between academic ability and FA, particularly for maths ability.

It is also possible that the environment may moderate the relationship between a child's brain architecture and their academic outcomes. For example, neurobiological risk factors such as low FA, might have more of an impact in lower SES participants as they may have reduced access to other factors that provide resilience in comparison to higher SES participants. Brito et al. found that family income and parent education moderated rather than mediated the relationship between cortical thickness and reading skills and vocabulary (Brito et al., 2017). Specifically, they found that the associations between higher cortical thickness and poorer outcomes was strongest for children from lower SES backgrounds. Gullick et al. also demonstrated this effect for FA; they found that parent education moderated the relationship between FA and reading skills using whole brain TBSS in children (Gullick et al., 2016). It would therefore be useful to conduct further research investigating whether the environment moderates the relationship between the FA connectome and academic outcomes.

Finally, it is important to note that differences in a child's brain associated with their environment are not necessarily evidence of a deficit due to lower SES. They may instead result from adaptive processes that provide resilience to children from disadvantaged backgrounds (Lipina and Evers, 2017). For example, Gullick et al. also found that the low- and high-SES groups had different patterns of association across the brain. Whilst reading ability was supported by high FA in some tracts common to both groups, the low SES group appeared to rely more on the visuospatial network than high-SES children. They suggest that this might be an adaptive mechanism to compensate for relatively under-stimulated areas elsewhere. As a result, a great deal of care is required in interpreting environmental differences in neuroimaging results. Further research into the complex associations between the environment, brain development and child outcomes will be useful in helping tease apart adaptive processes from neurobiological risk factors.

5.4.3 The structural connectome did not mediate the relationship between the environment and behaviour outcomes

FA did not mediate the relationship between the environment and behaviour outcomes. This is despite the fact that a several of the environment domains were related to behaviour outcomes and the FA connectome: child health, caregiver education, attitude

to child education and neighbourhood SES. The associations between FA strength and the environment and behaviour outcomes was more mixed than for the academic outcomes. Several connections had a negative outer weight in which high FA was related to lower environment scores. As FA covaried positively with the environment latent variable across almost every reliable connection, the negative associations indicate that high behaviour is related to lower FA in these tracts.

These results are echoed in a recent literature review of white matter and behaviour by Waller et al (Waller et al., 2017). Results were inconclusive across the 12 studies of white matter in children that met their inclusion criteria. In contrast to the adult studies, which demonstrated a consistent association between reduced FA and antisocial behaviour, there were null, positive and negative associations between FA and behaviour across the child studies. These differences were not accounted for by differences in sample type or age or ROI versus whole-brain methodology. We did not identify any studies that considered the mediating effect of FA on the environment-behaviour association.

It is possible that associations between FA and the behaviour outcomes might not be strong enough in our healthy sample to identify a significant effect. For example, Loe et al. found no significant association between FA and internalising or externalising behaviour (from the child behaviour checklist) in full-term children (Loe et al., 2013). Preterm children did have associations across the whole white matter skeleton and in several brain regions. Ikuta et al. found that higher BRIEF scores were associated with lower FA in young adults with autism in the cingulum bundle, but not in typical young adults (Ikuta et al., 2014). However, a large study by Decety et al. recruited 110 10-year-old children with a wide range of conduct disorder problems from severe to no problems. Despite the large differences in behaviour outcomes across children, no association between FA and the number of conduct disorder symptoms were found across the white matter skeleton using TBSS (Decety et al., 2015).

It might also be the case that other measures of brain structure are more sensitive predictors of behaviour outcomes in children. For example, Decety et al. did find significant associations between the number of conduct disorder symptoms and axial and radial diffusivity (Decety et al., 2015). Furthermore, despite inconsistent results during development, the association between FA and behaviour is well established by adulthood suggesting that FA might be important for behaviour but that a significant association takes longer to manifest and is not reliably related during development

(Waller et al., 2017). There may be other FA-based measures, such as the rate of development of FA, that might be more reliable predictors of child behaviour at this stage in their development. For example, a recent longitudinal study of 845 children found that global FA at the first time point did not predict changes in externalising or internalising behaviours (Muetzel et al., 2018). However, they found that higher behavioural difficulties at the first session predicted smaller increases in global FA over time. It would be interesting to investigate further markers of brain structure in our sample.

In summary, we have demonstrated the importance of whole-brain connectome methods and the application of PLS methods for capturing associations between brain structure, child environment and development across distributed brain systems. In particular, we have found that the association between a child's environment and their brain architecture is not localised to specific regions. Rather, whole-brain connectomics identified a set of widespread connections between brain regions across which the white matter structural integrity was significantly related to multiple domains in a child's environment. This has profound theoretical implications as to how the environment relates to brain development and sheds light as to why we might see such a variety of developmental outcomes related to a child's environment. To our knowledge, this is the first study to demonstrate an association between the environment and the FA connectome and that differences in white matter integrity might provide a mechanism underpinning the relationship between a child's environment and their academic ability.

6 ENVIRONMENT AND THE FUNCTIONAL CONNECTOME

6.1 Introduction

A surprising discovery in recent years has been that resting-state functional connectivity contains a wealth of information related to many different developmental outcomes. Defined as the correlation between activity in brain regions whilst a participant is at rest, resting-state networks have been related to reading (Hampson et al., 2006b; Koyama et al., 2011, 2013; Wang et al., 2012; Zhang et al., 2014; Zhou et al., 2015) and maths ability (Evans et al., 2015; Jolles et al., 2016b, 2016a; Price et al., 2018; Supekar et al., 2013), behaviour (Pu et al., 2017) and cognition (Astle et al., 2015; Barnes et al., 2015; Hampson et al., 2006a; Stevens et al., 2012). These individual differences in intrinsic functional connectivity are thought to reflect repeated co-activation between brain regions in the past (Guerra-Carrillo et al., 2014; van den Heuvel and Hulshoff Pol, 2010). Therefore, individual differences in the strength of a functional connection provide a measure of the efficiency of communication across that connection. As such, resting-state brain activity can be used as a marker of experience-dependant brain development, offering invaluable insights into the development of a child's functional brain architecture. As described in chapter 5, this efficient coordinated integration of distributed brain networks is fundamental to child outcomes. Thus, it is important to better understand how differences in a child's environment might translate into differences in the functional architecture of their brain and whether this mechanism underpins key childhood outcomes.

Resting-state connectomes have been found to be remarkably consistent within and between participants across different sessions and scans (Damoiseaux et al., 2006; Shehzad et al., 2009; van de Ven et al., 2004; Van Dijk et al., 2010). In fact, individual variability in functional connectomes has proven robust and reliable enough to accurately identify individuals from a large group (Finn et al., 2015). This functional connectome 'fingerprinting' is increasingly being utilised as a biomarker of typical and atypical brain development in children and adults. For example, not only have

distributed patterns of atypical activity in resting-state data been related to a number of developmental disorders (Barch, 2017) such as ADHD (Cao et al., 2014; Matthews and Fair, 2015), ASD (Hull et al., 2016; Jung et al., 2017; Matthews and Fair, 2015; Mevel et al., 2015; Nunes et al., 2018; Uddin et al., 2013; Vissers et al., 2012) and Schizophrenia (Andreou et al., 2014; Fornito et al., 2012; Schmidt et al., 2014; Tononi and Edelman, 2000), functional connectomes have even been used to classify individuals into disorder categories (Craddock et al., 2009; Du et al., 2018; Fair et al., 2012). In another example, Dosenbach et al. predicted individual brain maturity across development in typically developing 7-30 year olds (Dosenbach et al., 2010). Maturation was characterized by a weakening of shorter-range connections and strengthened long-range connections with age. Furthermore, there is evidence that changes in resting-state connectivity can be seen due to training. For example, improvements in working memory after training was associated with increased resting-state connectivity between MEG scans before and after training in a sample of 8-11 year old children (Astle et al., 2015). Similar training induced changes in connectivity have been found after training in cognitive skills (Cao et al., 2016; Mackey et al., 2013), mindfulness (Taren et al., 2015), motor skills , musical skills (Luo et al., 2014) and exercise (Voss et al., 2010).

Resting-state connectivity holds particular appeal in comparison to other neuroimaging modalities for a number of reasons. Primarily, in contrast to task-based functional imaging, it is independent of task design, participant strategy and level of engagement and training demands (Bokde et al., 2006; Greicius and Menon, 2004; Matthews and Fair, 2015; Rombouts and Scheltens, 2005; Tian et al., 2006; van den Heuvel and Hulshoff Pol, 2010). Resting-state protocols are easily replicated across studies and have been shown to be consistent across different scanners, field strengths, and statistical approaches (Biswal et al., 2010). Furthermore, it is easy to undertake in populations who might struggle with task-based imaging, such as young children or patient populations (Mevel et al., 2015) and stable estimates of connectivity can be found with acquisition times as short as 5 minutes (Van Dijk et al., 2010). In addition, whilst the structural and functional architecture of the brain are closely related (Hagmann et al., 2008; Hermundstad et al., 2013; Honey et al., 2009), there is evidence that differences in the functional connectome can be identified that are not observable at the level of structural connectivity (Pang et al., 2016). Indeed, the association between the structural and functional architecture of the brain is found to be weakest in children

(Hagmann et al., 2010; Supekar et al., 2010; Uddin et al., 2011), indicating that resting-state functional connectivity offers additional and complementary insights to those highlighted in chapter 5. Therefore, resting-state functional connectivity holds particular potential for investigating the relationships between the environment and brain development in children.

Task-related differences in functional connectivity have been related to SES using executive functioning, attention, language and reading, memory and emotion regulation tasks whilst participants underwent neuroimaging. Given this, it is likely that SES will also be associated with resting-state connectivity strength between regions across the brain. However, there has been limited research into the relationship between resting-state functional connectivity and a child's environment. Sripada et al. found that child poverty is associated with reduced default mode network connectivity in adults, using functional MRI (fMRI) (Sripada et al., 2014). Barch et al. also used resting-state fMRI to demonstrate that early childhood poverty was associated with reduced connectivity between the amygdala, hippocampus and several regions across the brain at age 12 (Barch et al., 2016). Recently, Chan et al. found that lower SES was associated with reduced segregation between brain regions using resting-state fMRI (Chan et al., 2018). A few studies have also investigated the relationship between resting-state EEG and SES measures, although findings appear to be mixed. Otero et al. found that lower SES children had higher delta and theta power in frontal and central regions and lower alpha power in occipital sensors than higher SES children, which they interpreted as demonstrating a maturational lag (Otero, 1997). Tomalski et al. found significantly lower frontal gamma power in infants, but no differences in alpha or theta power (Tomalski et al., 2013). Tomarkin et al. found that SES was related to frontal brain asymmetry (Tomarken et al., 2004). Finally, Brito et al. found no SES related differences in EEG at birth (Brito et al., 2016). Overall, these studies provide evidence that SES is generally related to connectivity between regions across the brain, but the small number of studies, mixed results and different approaches means that there is a great need for further research.

In addition, we could find only one study investigating how wider aspects of a participant's environment are associated with the resting-state connectome. Smith et al. used resting-state fMRI in a large adult sample to investigate the profile of demographic, environmental and behaviour variables that covaried most strongly with the functional connectivity across the brain (Smith et al., 2015). They identified a single

significant component using CCA (similar to PLS) and found that generally 'positive' traits were associated with higher functional connectivity and 'negative' traits were inversely associated. We wish to build on this research by investigating the relationship between multiple environmental domains and resting-state functional connectivity in children.

Furthermore, in chapter 5 we found that a child's structural brain architecture mediated the relationship between the environment and academic ability. In a similar way, given the relationship between resting-state connectivity and developmental outcomes, we hypothesize that environment-related individual differences in the functional connectome may underpin differences in child academic outcomes and behaviour. For example, a 'positive' childhood environment might facilitate better functional integration across a child's brain necessary for the development of academic skills and behaviour. This has been investigated in a task-based study by Noble et al. (Noble et al., 2006b). They found SES moderated the relationship between brain activity measured during an fMRI pseudoword reading task and phonological ability. However, we could find no papers investigating resting-state connectivity as a mechanism underlying associations between the environment and child outcomes. We will address this by investigating whether resting-state functional connectivity mediates the relationship between a child's environment and their developmental outcomes.

Finally, recent developments have made it possible to investigate relationships across the whole-brain resting-state functional connectome. As described in chapter 5, child outcomes require co-ordinated activity across distributed brain regions and, in line with the resting- and task-based functional neuroimaging studies available, individual differences in environment and developmental outcomes are likely to relate to connectivity strength in widespread connections across the brain. However, few studies have utilised whole-brain connectomics to study these relationships. For example, we could find only two studies investigating the relationship between the environment and the whole-brain functional connectome and these were in adult populations (Chan et al., 2018; Smith et al., 2015).

In particular, the use of MEG is rapidly gaining popularity as a neuroimaging modality for functional connectomics (Brookes et al., 2011; Pang et al., 2016). MEG measures the magnetic fields produced by electrical activity in the brain, making it possible to probe the electrophysiological basis of resting-state networks. To date, the majority of resting-state studies have focussed on the use of blood oxygenation level dependant

fMRI. In contrast to MEG, this is an indirect measure of brain activity related to haemodynamics and, as a result, has a relatively low temporal resolution. Furthermore, the research investigating the relationship between resting-state activity and SES using EEG has been conducted in the sensor space. Estimating the activity of regions in the brain in source space offers superior localization of functional connectivity in comparison to sensor space (Li et al., 2017). In addition, magnetic fields are not distorted by conductivity across the head making source localization more accurate than EEG. However, source localization is an ill-posed inverse problem in that there are far fewer sensors to estimate the local brain activity than 'sources' in the brain. As a result, 'source leakage', in which neighbouring regions in the brain are temporally correlated due to the source estimation, results in artefactual correlations between sources. This has greatly hampered the application of MEG to functional connectomics. A number of recent advances have been suggested to account for this (Baccalá and Sameshima, 2001; Brookes et al., 2011; Colclough et al., 2015; Hipp et al., 2012; Kaminski and Blinowska, 2014; Nolte et al., 2004; Stam et al., 2007). In particular, Colclough et al. have developed a symmetric multivariate leakage correction which removes zero-lag correlations across all ROI's in the brain (Colclough et al., 2015). In contrast to other source leakage corrections, this method has made it possible to conduct whole-brain functional connectome mapping using resting-state MEG data. This presents an exciting opportunity to utilise the temporal and spatial resolution of MEG to investigate the relationships between a child's environment and their resting-state functional brain activity using whole-brain connectomics.

In summary, PLS methods will be used to explore the associations between environmental variables and the resting-state functional connectome to answer three questions. First, what are the key aspects of a child's environment most strongly associated with their functional connectome? Second, does the wider environment mediate the relationship between the standard SES measures and their functional connectome? Finally, does the functional connectome act as a potential mediating mechanism underlying the relationship between the environment and academic and behaviour outcomes? To our knowledge, this is the first study to investigate these associations using the whole-brain functional connectome in children, and the only study to use resting-state MEG at any age.

6.2 Methods

6.2.1 Participants

9-minute resting-state scans were acquired for 95 children (2 children did not complete the MEG session) at the MRC Cognition and Brain Sciences Unit, Cambridge, U.K. A total of 12 scans were excluded: 5 due to scanner errors, 7 as they had less than 5 minutes of useable data (described in more detail below), and the 3 participants that were excluded due to having a large percentage of missing questionnaire data. The resulting sample included 80 children (38 males, mean age 10:0 and age range 7:0-12:9).

6.2.2 Resting-state MEG acquisition and pre-processing

MEG data were acquired using a high-density VectorView MEG system (Elekta-Neuromag) containing a magnetometer and two orthogonal planar gradiometers at 102 positions. Five head position indicator (HPI) coils were attached to the child's head in order to monitor the child's head movements throughout the recording and their position was recorded using a 3D digitizer (FASTRACK, Polhemus) in addition to over 50 additional points distributed over the scalp. Pulse was measured using an electrode attached to each wrist (ECG) and eye movements (EOG) were recorded using horizontal and vertical electrooculograms. Data were sampled at 1Khz. Small children were seated on a booster seat to ensure that the top of their head was in contact with the scanner. Children were monitored by video camera throughout the scan. During the 9-minute resting-state scan, children were asked to sit as still as possible, close their eyes and let their mind wander.

Data were first pre-processed using MaxFilter version 2.1 (Elekta-Neuromag) to remove external noise by signal space separation and to compensate for within-scan head movement at each time point using the HPI coils. The MaxFilter works by transforming the data to a set of virtual sensors. The following pre-processing steps were implemented using the OHBA Software Library (OSL v2.0.3) ("OHBA Analysis Group," 2017) and SPM12 (Penny et al., 2011; Wellcome Trust Centre for Neuroimaging, 2014). The data were downsampled to 250Hz. The sensor space time-courses for each subject at each sensor were visually inspected and any sections with large jumps were removed as this can affect the time-frequency analysis. Blinks, saccades and pulse-related artefacts were removed by running a temporal independent

component analysis (ICA) using fastICA on the sensor space time-courses. Components that correlated highly with the EEG and EOG time-courses (according to both Pearson correlation and visual inspection) were removed. Components dominated by 50Hz noise were also removed to reduce electrical interference noise.

6.2.3 MEG source reconstruction (beamformer)

Each subject's MEG data was co-registered to their T1-weigthed structural MRI image acquired using a 3T Siemens Tim Trio and an MPRAGE sequence. The standard MNI template was used for 7 children that did not undergo a T1-weighted MRI. The co-registration was performed using the digitized scalp locations and fiducial markers using an iterative closest point algorithm using SPM12. A forward model was fitted using a single shell homogeneous head shape model for each subject (Mosher et al., 1999). Data were initially band passed filtered to between 1-30Hz as it has been found that these slower frequencies are better for considering functional connectivity with MEG (Luckhoo et al., 2012). Finally, whole brain source-space activity was estimated for each point in an 8mm grid across the brain for each subject using a linearly constrained minimum variance beamformer (Van Veen et al., 1997). The reduced dimensionality of the data introduced by the signal-space separation algorithm was taken into account (Woolrich et al., 2011).This resulted in a set of estimated time-courses of brain activity for each child for source locations across the brain.

6.2.4 The functional connectome

The functional connectome was constructed by estimating the correlation in brain activity between the average time-courses for each pair of brain regions to construct connectivity matrices for each participant. This was conducted using the MEG-ROInets package developed by Colclough et al. (Colclough et al., 2015) using MATLAB 2012b (The MathWorks Inc., 2012), the FMRIB Software Library (FSL) (Jenkinson et al., 2012; Woolrich et al., 2009) and FieldTrip (Oostenveld et al., 2011).

68 ROI's (34 in each hemisphere) were defined using the Desikan-Killiany parcellation of the MNI template (Desikan et al., 2006). The subcortical regions used in the structural connectome were excluded due to the difficulty in detecting subcortical sources using MEG (Hillebrand and Barnes, 2002). For this analysis, we focussed on activity in the beta band because pervious research has found that resting-state connectivity is typically strongest in this frequency range (Barnes et al., 2015; Brookes

et al., 2011, 2016). The source time-courses were bandpass filtered to 13-30Hz. In order to obtain a single time-course for each ROI, the time-courses were normalised to have a positive peak height of unity and a PCA analysis was conducted for each set of sources in an ROI. The primary principle component coefficients were then used to calculate a single time-course for each ROI that best represented the majority of the variance across the ROI. Next, symmetric multivariate orthogonalization was conducted to correct for source leakage with zero temporal lag, introduced by the source-reconstruction. Firstly, the set of ROI time-courses were orthonormalized to remove any correlation between them that had zero time lag. Secondly, the amplitudes and orientations of the ROI vectors were iteratively adjusted until the corrected time-courses were as close as possible to the original time-courses. Finally, the absolute signal amplitude for each ROI at each time-point was estimated using the Hilbert transform and the data was low-pass filtered to 0.5Hz and downsampled to 1Hz to focus on the low frequency power fluctuations that have been shown to demonstrate functional connectivity using MEG (Brookes et al., 2011; Colclough et al., 2015)

Connectivity matrices were calculated for each child by finding the partial correlations between each pair of ROIs. The partial correlation, in which the effect of all other ROI's was removed, was selected as this emphasises direct network connections rather than indirect connections and has been found to be better at defining true network structure in comparison to full correlation (Colclough et al., 2016, 2015; Marrelec et al., 2006; Smith et al., 2011). It has also been shown that applying regularisation that encourages sparsity to supress null edges reduces noise and improves the connectome estimation from partial correlations (Colclough et al., 2015; Varoquaux and Craddock, 2013). According to common practice, we applied L1 regularisation using the graphical LASSO (Friedman et al., 2008). The strength of the regularisation was chosen using 10-fold within-subject cross-validation, in order to select the regularisation that minimised the Akaike information criterion, as applied in Colclough et al., 2015. This resulted in a sparse partial correlation connectivity matrix for each participant in which each data point represents an estimate of the strength of functional connectivity between each pair of regions in the brain (see Figure 50). Finally, the top triangles of each matrix for each subject were reshaped into single vectors and combined across subjects, resulting in a 2278 X 80 dataset in which each column represents the FA between each pair of regions for each of the 80 participants.

Figure 50 Group-average partial correlation matrix indicating the ROI-by-ROI connection strength estimated from resting-state MEG. The ROIs are grouped by lobe and hemisphere (lh= left hemisphere and rh= right hemisphere).

6.2.5 PLS between the environment and the functional connectome

PLS was used to identify the profile of environmental factors that most strongly covaries with the functional connectome. Each variable was standardised and the PLS analysis was conducted as in section 3.2.2. The environmental variables and connectome data used as dataset X and Y respectively. Rubin's rules were used to pool the PLS results across imputation and the significance and reliability of the PLS model were assessed using permutation (N permutations = 1000) and bootstrap (N bootstrap samples = 500) analyses.

6.2.6 PLS mediation

We applied 3-block Canonical PLS to the standardised datasets as detailed in Section 3.2.3 to investigate mediation effects. First, the mediating effect of the wider environment on the SES-connectome relationship was investigated using the three standard SES measures, equivalised income, caregiver education and occupation as dataset X, the other environmental domains as dataset M and the connectome data was used as dataset Y. Primary caregiver's hours work per week was excluded from the wider environment dataset.

Secondly, to investigate whether a child's resting-state connectivity mediated the relationship between a child's environment and their academic and behaviour outcomes, the full set of environmental domains were used as dataset X, the connectome was used as dataset M and either the two academic variables or two behaviour variables were used as dataset Y.

6.3 Results

6.3.1 PLS: What is the environmental profile that most strongly relates to a child's functional connectome?

Two-block PLS was conducted using the environmental and functional connectome datasets. The first pair of latent variables did not significantly explain the covariance between the environmental domains and the functional connectome (p = 0.098, explained variance: 18% environment and 2% functional connectome, average correlation between environment and functional connectome latent variable = 0.90 Figure 51).

Bootstrapping methods allow us to identify the variables that are consistently found to covary most strongly with the latent variable of the other dataset. Bootstrapping the data indicates that there is a mixed pattern of covariance across the brain. The strongest 25% (absolute value) of connections are plotted for visualisation in Figure 52. In just over half of the connections that were identified as reliable, lower partial correlation was related to higher environment scores. Higher caregiver education, occupation, child health and equivalised income and lower technology use and discipline were reliably found to relate to the functional connectome latent variable (see Figure 51). However, as the overall covariance between the latent variables is not found to be significant in

comparison to permuted data, we would be very cautious in assigning relevance to the bootstrap results.

Figure 51 The environment and the functional connectome: outer weights for the first latent variables, averaged across imputations

Figure 52 The environment and the functional connectome: a topographic and circle plot of the strongest 25% of edges that were reliably related to the environmental domains based on the bootstrap of the first PLS component, coloured by the sign of the loading onto the first component (orange, positive; blue, negative). Edges are grouped by brain hemisphere (L/R) and lobe (given in the key).

6.3.2 PLS Mediation

6.3.2.1 Does the wider environment mediate the relationship between standard SES measures and the functional connectome?

The wider environment significantly mediated the relationship between SES and the functional connectome for the first latent variable (indirect path coefficient ab=0.42±0.07, 95% CI [0.31,0.58]). The path model for the first component can be seen in Figure 53 and the variable weights can be seen in Figure 54. Note that the direct path between SES measures and the functional connectome remained significant even when the other environmental domains were accounted for, indicating that the SES measures explained unique variance in the functional connectome that was not related to the

wider environment domains measured. The second component was significant and the third was marginally significant.

Each of the SES measures loaded strongly onto the first latent variable. Inverse technology use, child health, other language use, primary caregiver skills, inverse discipline and attitude to child education loaded reliably onto the mediating latent variable. Bootstrapping the data indicates that there is a mixed pattern of covariance across the brain, similar to the pairwise relationship in 6.3.1. The strongest 25% (absolute value) of connections are plotted for visualisation in Figure 55.
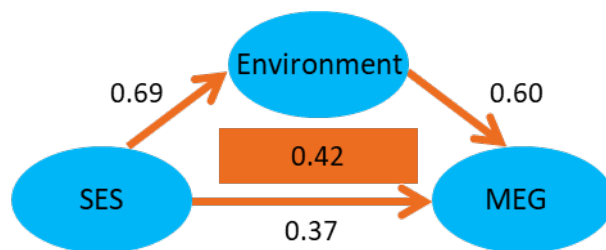
Figure 53 PLS mediation analysis between SES, the wider environment and the functional connectome: the path model
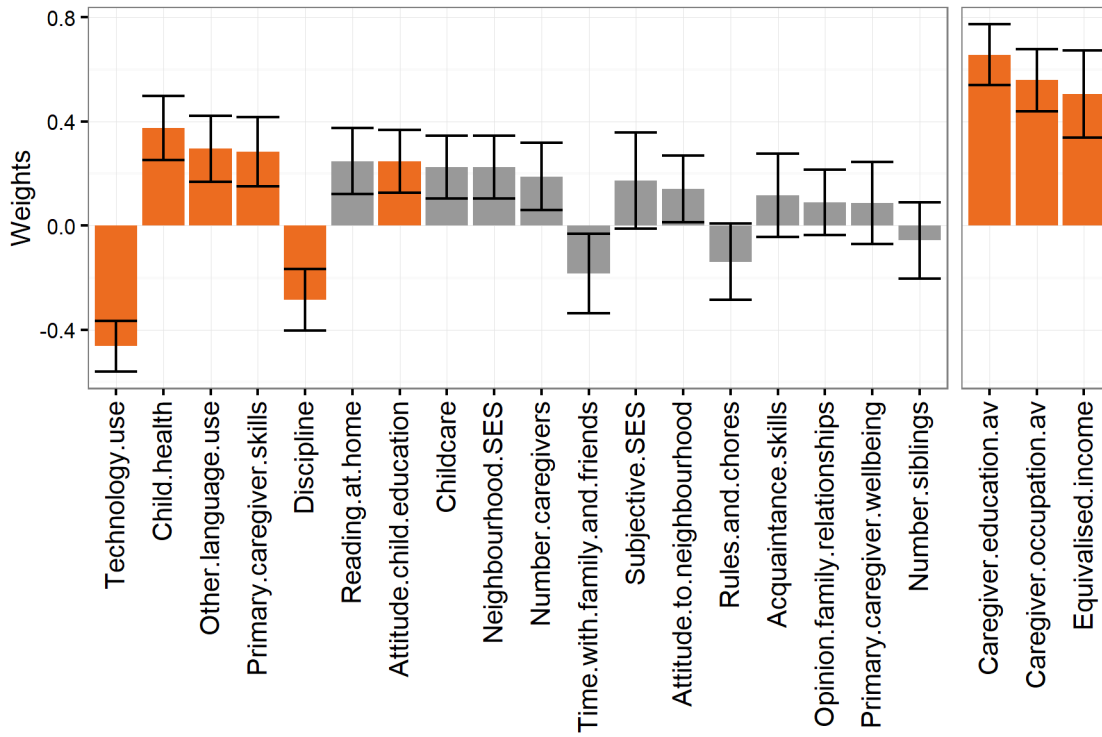
Figure 54 PLS mediation analysis between SES, the wider environment and the functional connectome outer weights for the first latent variables, averaged across imputations
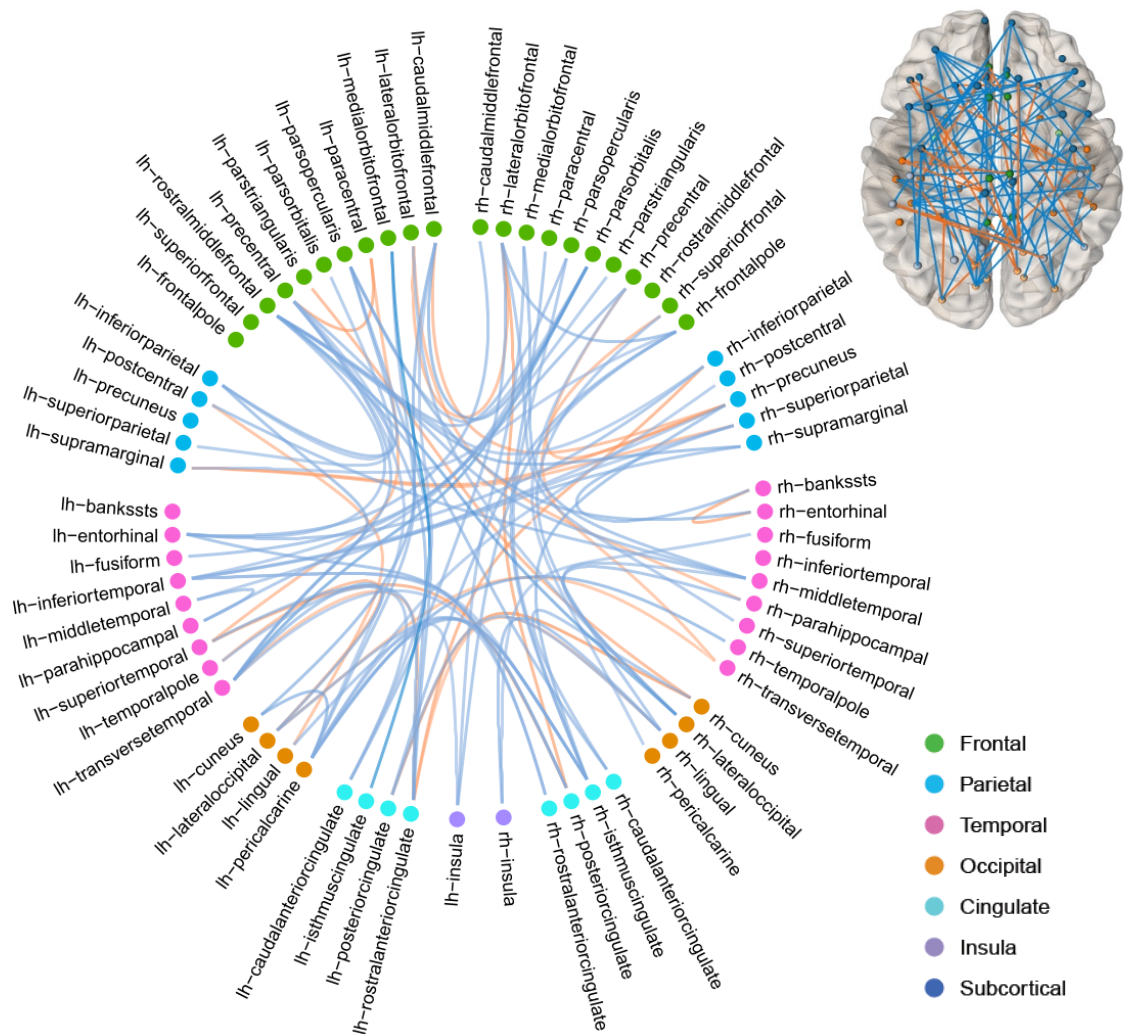
Figure 55 PLS mediation analysis between SES, the wider environment and the functional connectome: a topographic and circle plot of the strongest 25% of edges that were reliably related to the latent variable

6.3.2.2 Does the functional connectome mediate the relationship between a child's environment and academic ability?

The functional connectome fully mediated the relationship between the environment and academic outcomes in the first component (indirect path coefficient ab=0.93±0.17, 95% CI [0.54,1.21]). The path model can be seen in Figure 56 and the variable weights can be seen in Figure 57 and Figure 58. Maths and reading loaded equally strongly onto the academic latent variable. Inverse technology use, caregiver education, occupation, reading at home, child health, attitude to child education each loaded reliably onto the environment latent variable. There is a mixed pattern of covariance across the brain.
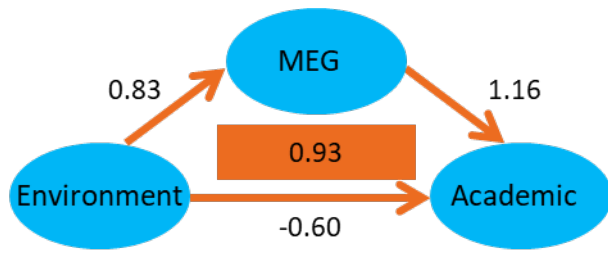
Figure 56 PLS mediation analysis between the environment, the functional connectome and academic ability: the path model
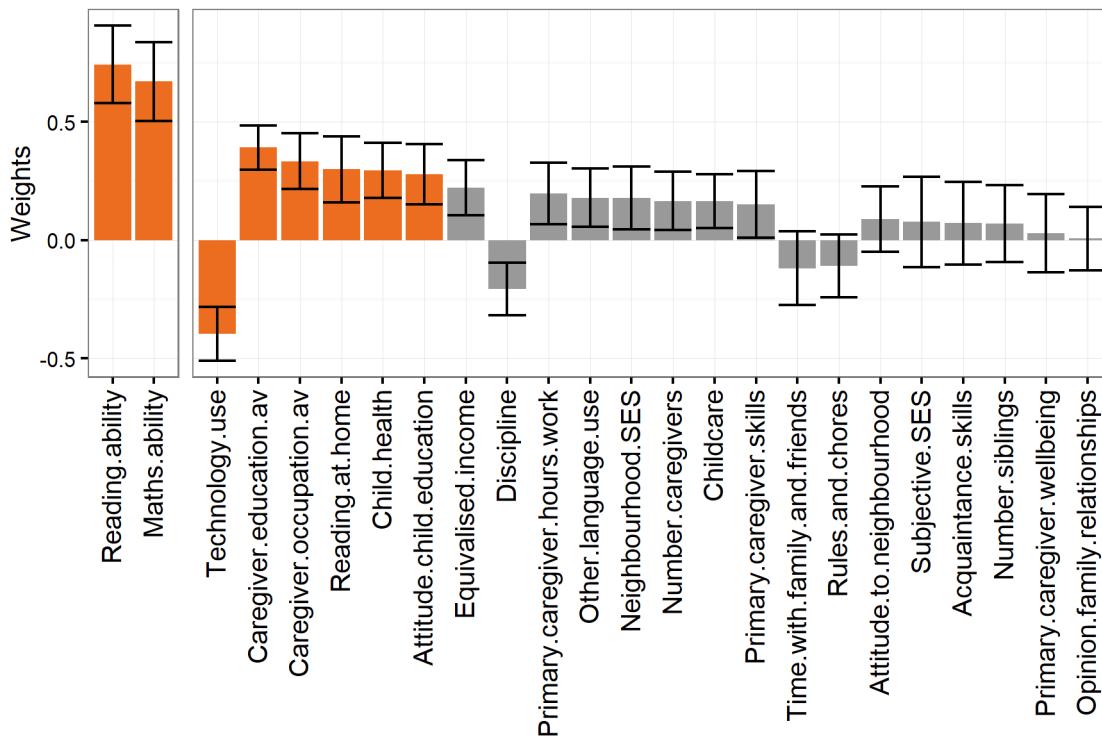
Figure 57 PLS mediation analysis between the environment, the functional connectome and academic ability: outer weights for the first latent variables, averaged across imputations
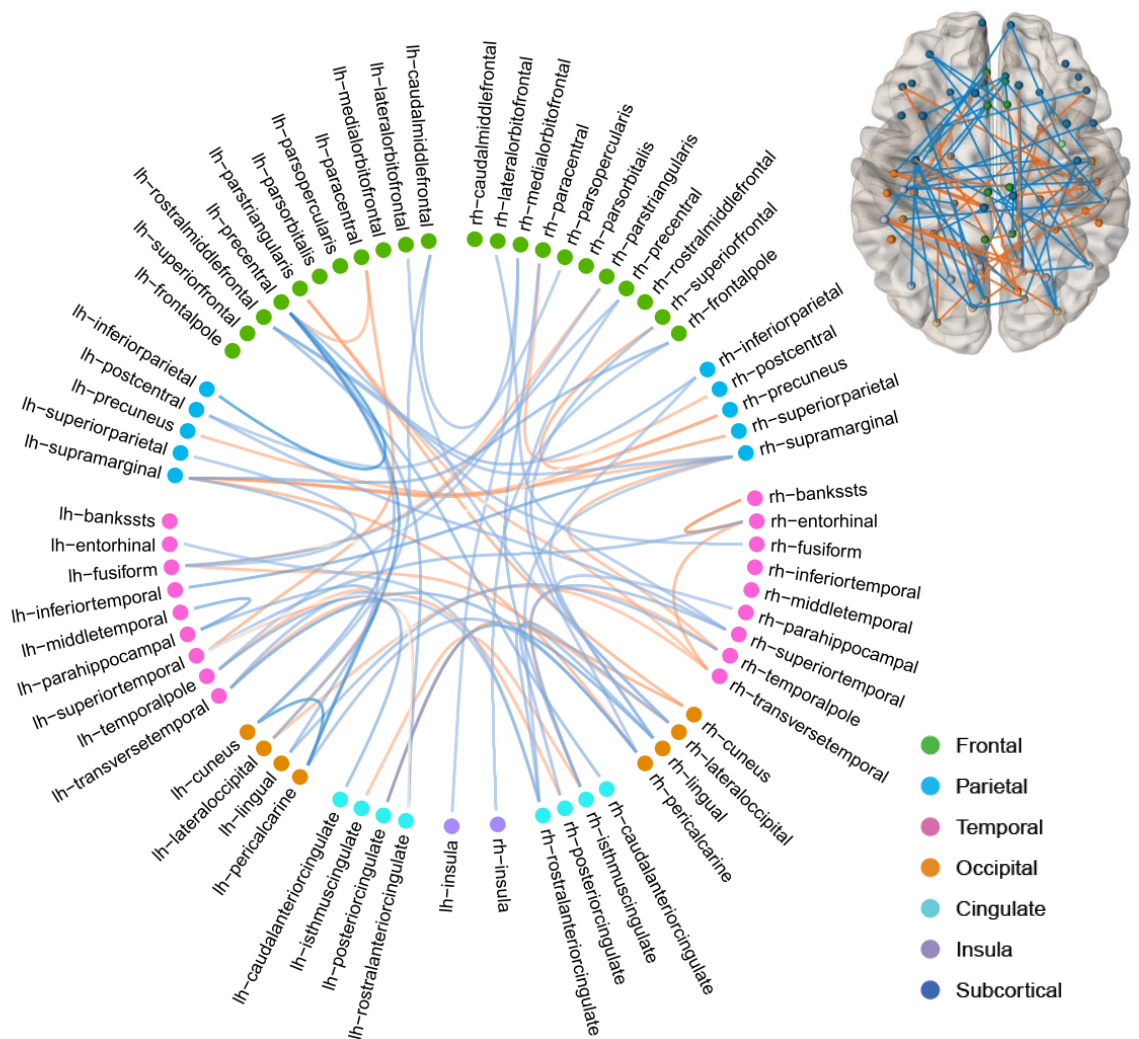
Figure 58 PLS mediation analysis between the environment, the functional connectome and academic ability: a topographic and circle plot of the strongest 25% of edges that were reliably related to the latent variable

6.3.2.3 Does the functional connectome mediate the relationship between a child's environment and behaviour?

The functional connectome fully mediated the relationship between the environment and behaviour outcomes in the first component (indirect path coefficient ab=0.73±0.22, 95% CI [0.22,1.10]). The path model can be seen in Figure 59 and the variable weights can be seen in Figure 60 and Figure 61. BRIEF and SDQ scores load equally strongly onto the behaviour latent variable. Inverse discipline, child health, inverse technology use, caregiver education, primary caregiver skills, reading at home, caregiver occupation, equivalised income, neighbourhood SES load onto the environment latent variable. There is a mixed pattern of covariance across the brain.
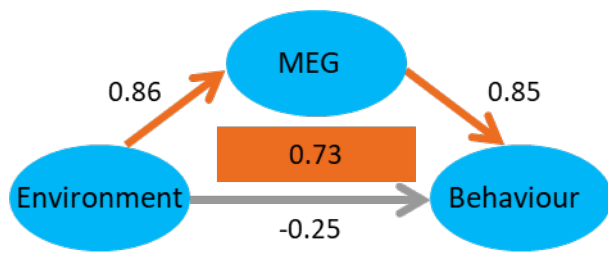
Figure 59 PLS mediation analysis between the environment, the functional connectome and behaviour: the path model
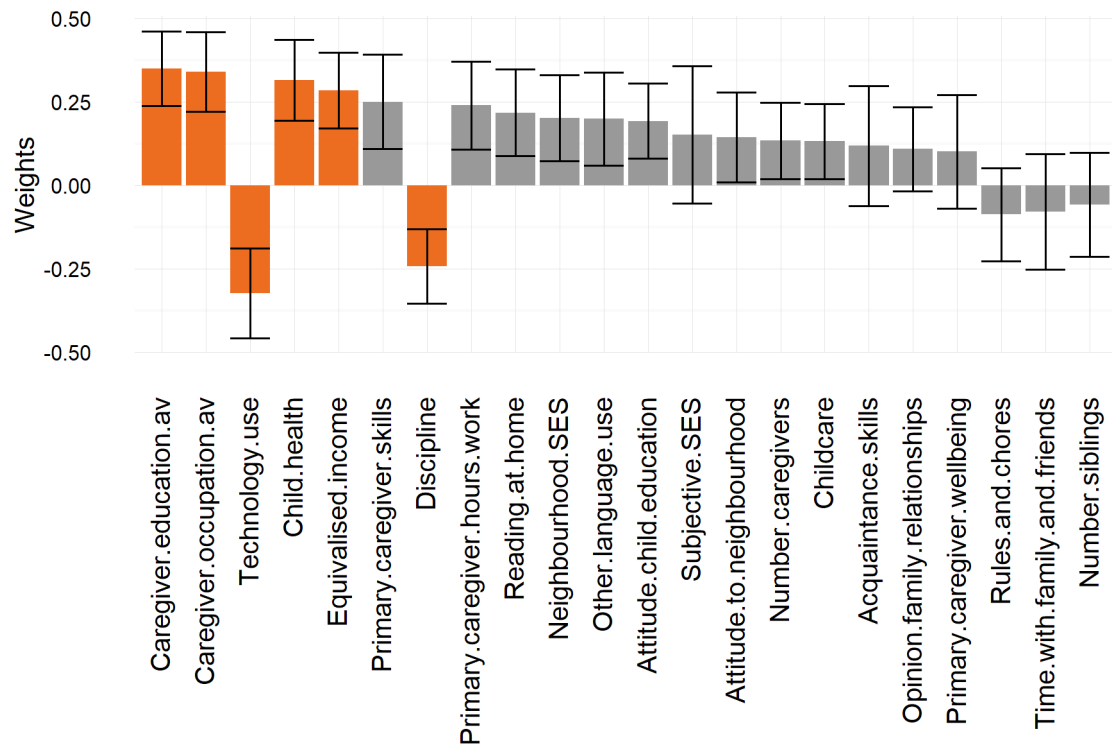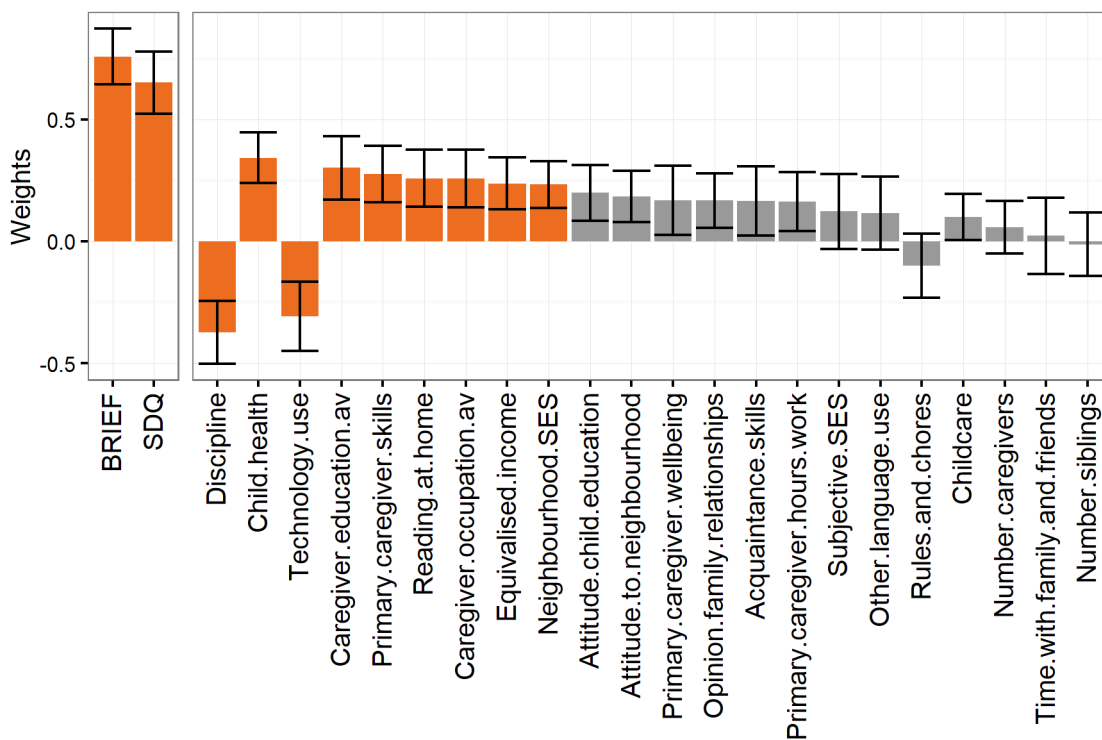

Figure 60 PLS mediation analysis between the environment, the functional connectome and behaviour: outer weights for the first latent variables, averaged across imputations
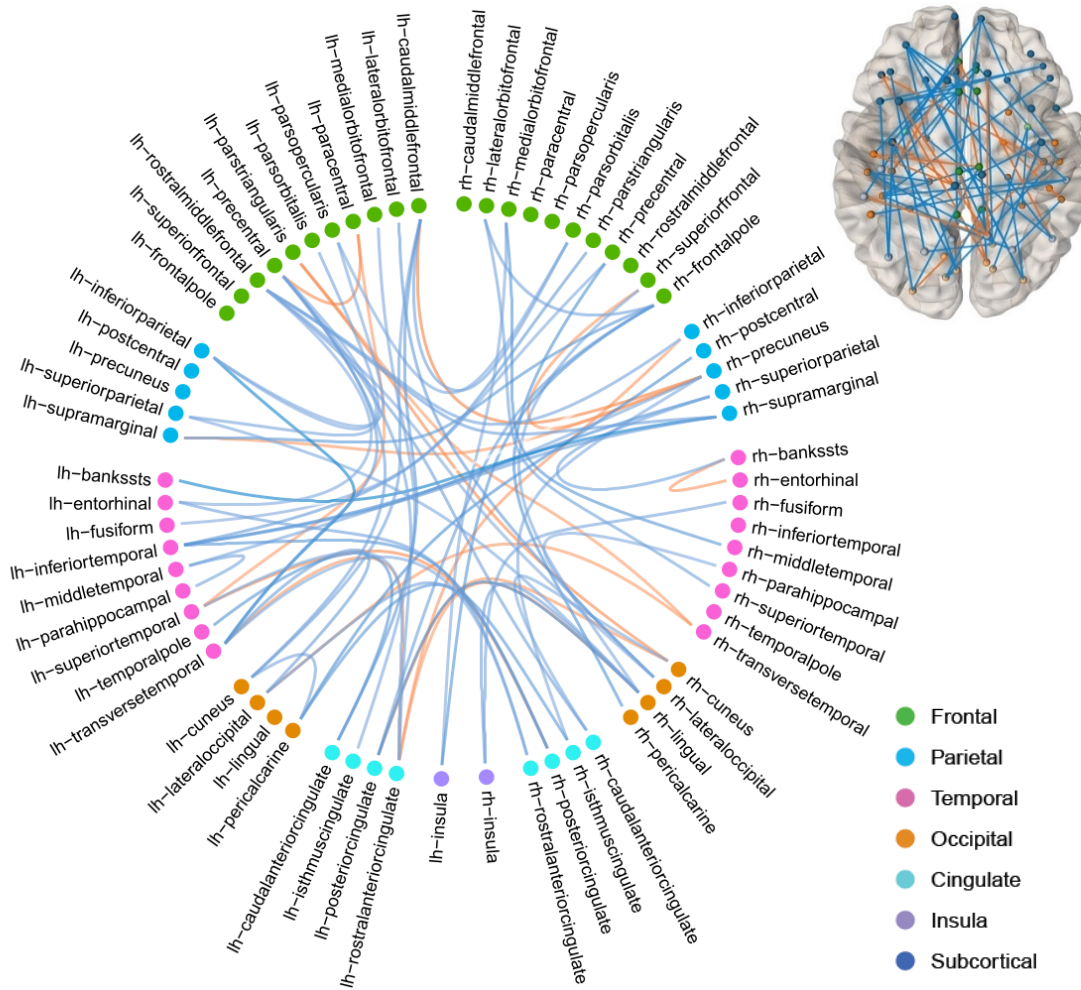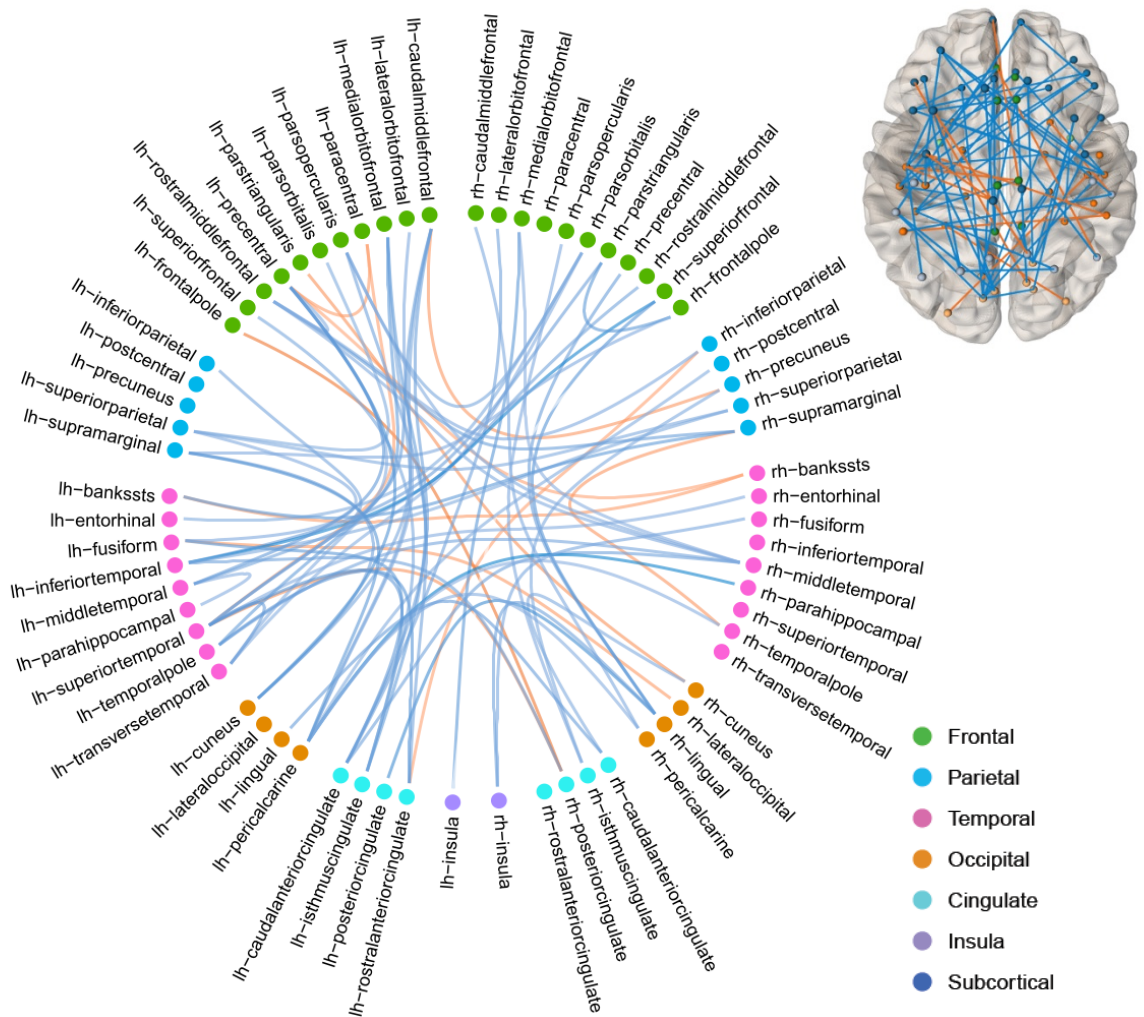
Figure 61 PLS mediation analysis between the environment, the functional connectome and behaviour: a topographic and circle plot of the strongest 25% of edges that were reliably related to the latent variable

6.3.3 PLS: The need for sparsity

We found that there was no significant pairwise relationship between the environment and the MEG connectome. In contrast, each correlation coefficient and path coefficient between latent variables was remarkably high. A few of the coefficients are greater than 1 due to the high collinearity between latent variables. This frames a key issue: if there are a large number of noisy variables, there are many ways that they can be combined to create a latent variable and so the PLS model can form latent variables that are highly correlated with each other, even when modelling noise. For example, latent variables from the permuted dataset are also highly correlated and so the PLS model is not found to be significant. In practice, this may result in false negatives. On the other hand, when

assessing the reliability of the mediation effects, the high correlations between latent variables result in almost all paths being reliable and the mediation effect is found to be significant. This, in turn, may result in false positives.

The suitability of PLS for a large number of variables in the context of comparatively small sample size has been called into question. For example, Nadler and Coifman demonstrate that the root mean squared error (RMSE) in the PLS model includes an error term that depends on $\sigma^2 \frac{p^2}{n^2}$ where σ is the noise level per variable, p is the number of variables and n is the sample size (Nadler and Coifman, 2005). Not only is resting-state MEG data particularly noisy, in the case of our connectome, $p \gg n$. As a result, this error term is likely to dominate the results. In addition, it has been demonstrated that a large number of irrelevant connections attenuates the outer weights considerably, making estimates inconsistent (Chun and Keleş, 2010).

This is a considerable issue in today's context of big data, and for neuroimaging applications in particular, where the sample size is typically small. A number of solutions have been proposed to address this issue, by performing variable selection to reduce p. There are two main methods to produce this sparsity: variable selection can either be performed on the dataset before PLS modelling or it can be integrated into the dimension reduction step within PLS, often known as sparse PLS. In the following sections, we illustrate these two methods using our data.

6.3.3.1 Imposing sparsity on the MEG connectome before PLS

There are many ways to impose sparsity on the MEG connectome. For example, LASSO regularization is used on our data to reduce noise in the partial connectome. However, the regularization is applied to each individual connectome separately. As a result, different connections are set to zero for each participant and so, for our sample, all connections have at least a proportion of children that have non-zero correlations. It would therefore be useful to impose a common sparsity across each of the participant connectomes. We tested this using multiple different methods and modelled the relationship between the resulting MEG datasets and the environment domains using PLS. The variable selection and PLS modelling were applied to each imputed dataset in turn and the p values were pooled as before. These included:

- Creating a mask by thresholding the group connectome: connections with z-value below a given threshold in the group connectome are set to zero for each participant

- Thresholding the individual connectomes based on various values, with and without additional consensus thresholding of 60%
- Using the 265 white matter connections from the MRI connectome to create a mask for the MEG connectome

None of these methods produced sparse connectomes that were found to significantly relate to the environment latent variable. For example, Figure 62 gives the average covariance between the connectome and environment dataset for a range of z-value thresholds applied to the group connectome (higher z value equates to greater sparsity as more connections are set to zero). The grey area is the covariance that the data must reach in order to be significantly better than the permuted data, averaged across imputations. The relationship is therefore not significant at any degree of sparsity imposed.
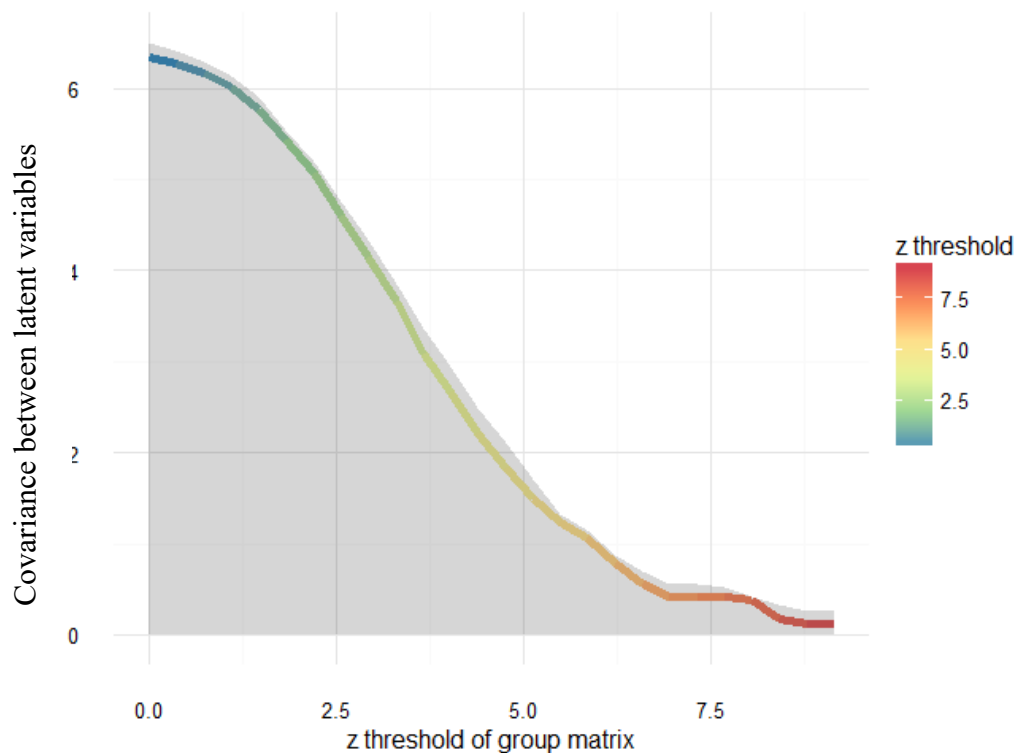
Figure 62 Average covariance between the environment and the functional connectome latent variables for a range of z-value thresholds applied to the group functional connectome. The grey area is the covariance that the data must reach in order to be significantly better than the permuted data, averaged across imputations

In addition, it is common to reduce the number of variables by conducting dimension reduction techniques on the MEG connectome. For example, we used PCA to summarize the variance in the connectome. When sets of the top components are used

as the connectome dataset, the PLS model is not significant for any number of principal components.

One possible reason for these results is that the variable selection and dimension reduction were conducted without taking into account the environmental data. As a result, the selected variables or principal components are not necessarily the ones that most strongly relate to the environmental domains. It would therefore be useful to use the environment domains to influence the variable selection step in order to select the most relevant variables for the model.

6.3.3.2 Sparse PLS: Imposing sparsity in the PLS model

In recent years, methods that integrate variable selection into the PLS algorithm have been developed to address the issue of high dimensionality. Predominantly, methods are based on LASSO regularisation of the weight vector during each loop of the algorithm (step D in section 2.3.1), in the form of an additional constraint (Chun and Keleş, 2010; Rohart et al., 2017). The equation to be maximised becomes:

$$max \sum_{j,k=1, j \neq k}^{J} c_{jk} cov(L_j, L_k) \ constrained \ by: \|a_j\|^2 = 1 \ and \ |a_j| < \lambda \ for \ j = 1, 2 \ldots J$$

Where λ is the penalization parameter. Lower values of λ result in more zero's in the outer weight vector and thus a greater degree of sparsity. Note that some studies have found this to not be sufficiently sparse (Chun and Keleş, 2010). The MixOmics package, used in this thesis, suggests an adaptation in which the number of variables to be selected, 'keepX', is used as the penalization parameter. The outer weights are estimated in each iteration and ordered by strength. The highest outer weights are selected up to the number of 'keepX' and the algorithm is repeated. 'keepX' can be different for each component and dataset. As the outer weight relates to the covariance between a given variable and the other latent variables, this ensures that a subset of variables are selected that are most relevant to explaining the relationship between datasets.

There are two related issues to address: selecting the sparsity strength and assessing significance. We might, for example, repeat the PLS analysis with a range of sparsity strengths, as proposed by Witten and Tibshirani (Witten and Tibshirani, 2009). At each value of 'keepX', the covariance between the resulting latent variables is compared to the covariances produced under permutation of the data. However, when we apply this

method to our data, we find that the results are not significant for any of keepX = 1, 5, 10, 25, 50, 75 and 100% of the data. Note that there is no precedence for using multiple imputation with sPLS. In this case, the sPLS model was fitted to each imputed dataset and p values were pooled as before. Figure 63 demonstrates that the (uncorrected) p values remain non-significant at each value of keepX.
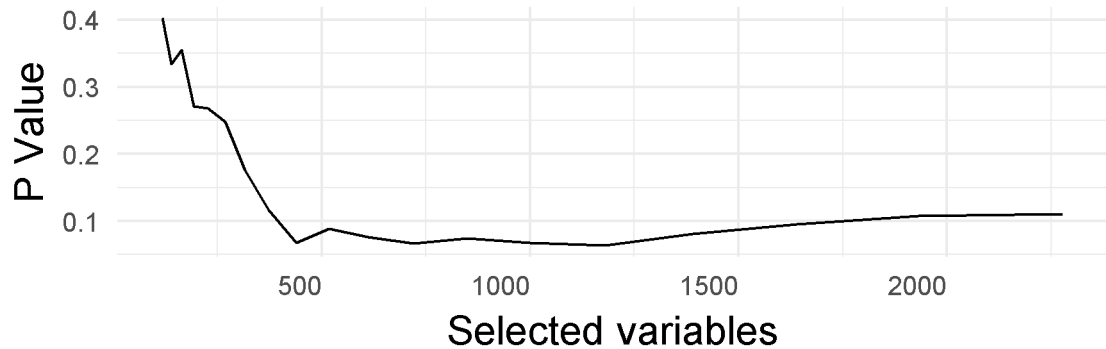
Figure 63 Uncorrected p-values for the covariance between the environment and the functional connectome latent variables for a range of 'keep X'.

One particular issue with this method is that the large amount of irrelevant data means that the permuted datasets are equally good at finding a set of 'keepX' MEG connections that combine to create latent variables that covary strongly. Le Floch et al. address this issue by suggesting that the data should first be filtered univariately based on the relationship between the variables in one dataset to the variables in the other dataset before the sPLS model is fitted (Le Floch et al., 2012). For example, they found that the correlation between a set of genetic polymorphisms and a functional brain network was only significant when they first filtered the Single Nucleotide Polymorphisms (SNPs). They calculated the p-values for the linear regressions between each genetic variable onto each of the 34 neuroimaging variables and ranked the genetic variables based on their lowest p value. Using this, they selected the top set of SNP's that were most strongly related to one or more of the neuroimaging variables, before submitting these to sPLS. They found that the model was only significant when the top 1000 SNPs were chosen before sPLS.

In order to test this on the functional connectome data, the covariance between every MEG connection and environment domain was calculated and the MEG connections were ranked based on the maximum covariance they had across the environmental domains. sPLS using keepX = 1, 5, 10, 25, 50, 75 and 100% of the data was fitted to the datasets using different filters by including only the top set of MEG connections up to a

given number. This was tested with filters of 100, 500, 1000 and all data and the resulting correlation and uncorrected p-values can be seen in Table 6. The covariance between pairs of latent variables has been converted to correlation for ease of interpretation. No combination of filtering and sparsity strength was significant.

Table 6 Uncorrected p-values and average latent variable correlation for a range of sPLS models with univariate filtering of 100, 500, 1000 and all data and keepX = 1, 5, 10, 25, 50, 75 and 100% of the data.

P values							
k/s	1%	5%	10%	25%	50%	75%	100%
100	0.428041	0.568759	0.594006	0.683206	0.535996	0.510468	0.603565
500	0.750888	0.739842	0.441999	0.650419	0.485925	0.529825	0.555976
1000	0.594024	0.393302	0.202269	0.389091	0.323252	0.387669	0.415229
All	0.573258	0.401522	0.290205	0.078611	0.06166	0.095846	0.093574

Average correlation							
k/s	1%	5%	10%	25%	50%	75%	100%
100	0.483406	0.613107	0.700958	0.793132	0.816155	0.806921	0.790335
500	0.655156	0.808121	0.85646	0.882894	0.893244	0.899591	0.899733
1000	0.702826	0.837058	0.874129	0.892413	0.907036	0.913842	0.916555
All	0.750085	0.867455	0.881474	0.888156	0.894143	0.89931	0.902165

It is clear that overfitting is still a problem despite the filtering and sparsity. The industry standard for dealing with this issue of overfitting is to use cross-validation methods. In the case of sPLS, the model is trained on a subset of the data (the training set) and then the outer weight vectors from this are projected onto a test dataset for a range of different sparsity strengths to calculate the test set latent variables. Typically, this is repeated with several folds of the data. For example, Le Floch et al. performed 10-fold cross-validation in which univariate filtering and sPLS models were fitted to 90% of the data and the model was projected onto the remaining 10% of unseen data to calculate the correlation between the latent variables in the test data (Figure 64). The process is repeated 10 times with a different test set each time and the average

correlation calculated. If the sPLS model has overfitted the training dataset, then the correlation will be low. Alternatively, the correlation will be high if the sPLS model generalises well to unseen data. Significance is then tested by shuffling the rows in one dataset and repeating the 10-fold cross-validation using this set several times to build up a null distribution of test set correlation.

For each sPLS parameter:
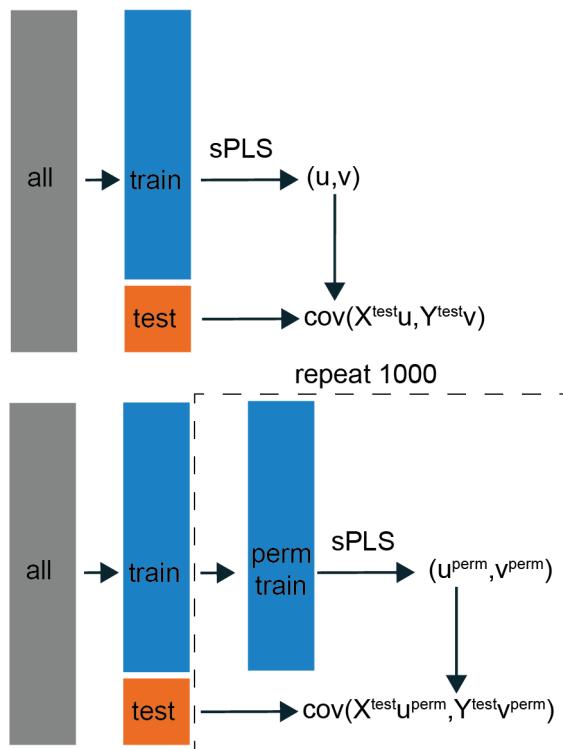
Figure 64 sPLS with 10-fold cross-validation method

However, a great deal of caution must be applied in using cross-validation techniques on our dataset. This is primarily due to the low sample sizes and the high degree of noise in MEG connectome data. For example, if we apply 10-fold cross-validation, then the test set only consists of 8 participants. Not only does the high degree of variability between participants mean that the testing set covariance is likely to vary wildly, covariance estimations are highly unstable at such small sample sizes. If we apply Le Floch's method to our data, we find that the covariance between latent variables varies strongly between the 10 different training and test datasets, with several covariances falling below zero in the test data. The test set correlation and corresponding p values from permutation testing for a range of filtering and sparsity strengths can be seen in Table 7. The average test set covariance is typically very low, and no model is significantly better than the permuted data.

Table 7 Testing set uncorrected p-values and average correlation between the environment and the functional connectome latent variables for a range of sPLS models with univariate filtering of 100, 500, 1000 and all data and keepX = 1, 5, 10, 25, 50, 75 and 100% of the data using 10-fold cross-validation.

P values							
k/s	1%	5%	10%	25%	50%	75%	100%
100	0.290602	0.6114	0.788656	0.769889	0.856027	0.587605	0.799548
500	0.25071	0.425845	0.53077	0.765254	0.848076	0.555256	0.850186
1000	0.104429	0.709547	0.470357	0.374376	0.644586	0.896281	0.775248
All	0.525246	0.293594	0.198438	0.12026	0.063418	0.060101	0.298767

Average correlation							
k/s	1%	5%	10%	25%	50%	75%	100%
100	0.180827	-0.03955	-0.17128	-0.12129	-0.06277	0.00248	-0.09982
500	0.105636	0.012957	-0.08147	-0.11914	-0.16277	-0.03923	-0.23683
1000	0.234622	-0.17194	0.052448	0.077097	-0.04325	-0.21769	-0.17131
All	-0.00097	0.125303	0.155318	0.162881	0.200092	0.173101	0.089893

It might be possible to improve the covariance estimation in the test data by increasing both the size of the test dataset and the number of times the data is divided into training and test datasets. This is possible if the data is randomly split into training and testing sets several times (e.g. 100 times) to build up a distribution of test set latent variable covariances. Unfortunately, if the permutation analysis was applied to each of these folds, the computation is typically prohibitively intensive for large datasets. Monteiro et al. addressed these issues for neuroimaging datasets by proposing a hyper-parameter optimisation method for sPLS in which 90% of the data is used to identify the best sparsity parameter and 10% of the data is retained as a holdout dataset to use for assessing significance using this sparsity strength (Monteiro et al., 2016). In this, 90% of the data is randomly split 100 times into a training set (80%) and testing set (20%) and the test set latent variable covariance is found for each split and averaged over 100

folds for each value of the sparsity parameter (Figure 65). The sparsity parameter that gives the highest average covariance is selected and the full 90% dataset is fitted using sPLS with this sparsity parameter and projected onto the 10% hold out set to calculate the hold out covariance. The 90% set is permuted 1000 times and the sPLS repeated to build up a null distribution for the hold-out latent variable covariance. P values for the hold-out covariance are then computed in the normal way. The entire process is then repeated 10 times using a different 10% hold-out set each time to calculate 10 p-values.

A) Use 90% of the data to choose sPLS parameters that have highest covariance
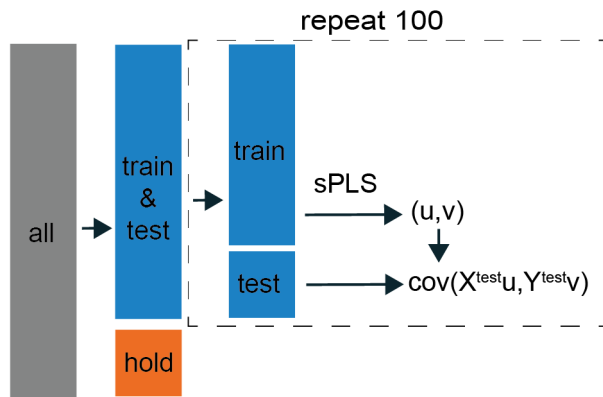
B) Use these sPLS parameters to test significance using the hold out set
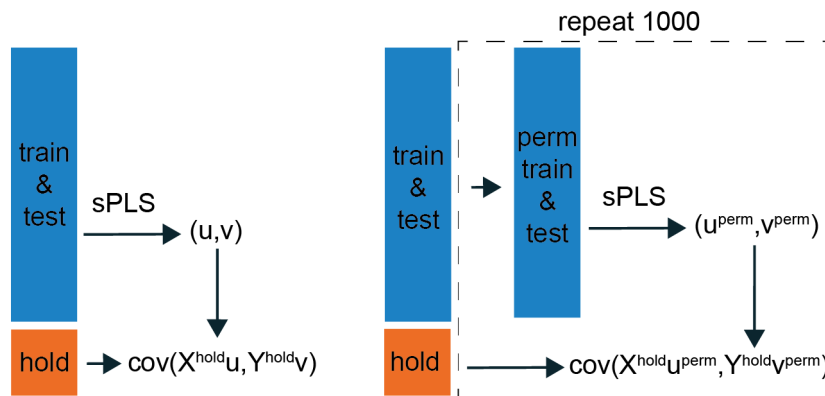
Figure 65 Hyper-parameter optimisation method for sPLS

However, this method still relies on 10-fold cross-validation for significance testing, which, in our sample, is too small a subset for stable results. In addition, the high degree of variance in the MEG connectome dataset results in a very wide distribution of latent variable covariance across the 100 random splits. This results in average covariances close to zero and the number of variables selected is therefore unstable and varies wildly between the 10 cross-validation folds. Furthermore, none of the folds were found to be significant.

6.4 Discussion

Initial research has highlighted the importance of functional connectivity as a key marker of child development, yet few studies have extended this to consider the potential impact of a child's environment. PLS methods were used to investigate the complex relationships between a child's environment and their whole-brain resting-state MEG connectome. However, high dimensionality and noise in the functional connectome presents a key challenge in this analysis. This is illustrated by the covariance and correlation between latent variables which was found to be very high, even under permutation. This results in the potential for false negatives in permutation tests and false positives using bootstrapping methods. This highlights the need for particular caution in the application of PLS methods when the number of variables is much greater than the number of participants due to overfitting of the PLS model. Note, that this is typically the case in neuroimaging studies due to the prohibitive cost and time required to collect data. As datasets become ever larger with the improvement of data collection, storage and processing methods, this challenge is increasingly being faced, such as in whole-brain analyses. As a result, it is imperative that variable selection methods are developed and applied to reduce the number of variables in the model.

Several methods of imposing sparsity were discussed and applied to the connectome dataset in section 6.3.3. In line with previous research, we found that imposing sparsity on the connectome data without taking into account the relationship to the other datasets, did not result in a significant PLS model between the environment and the connectome (Le Floch et al., 2012). For example, we did this by thresholding the connectome matrices based on correlation strength or applying a dimension reduction technique such as PCA– these methods were ineffective. Sparse PLS has been proposed for high dimensional datasets as a method that integrates variable selection into PLS algorithm. This is particularly appealing as it ensures that the variables selected are those most strongly related to the other datasets. We developed a means of implementing this alongside the multiple imputation approach. However, overfitting of the data remained a significant challenge at a range of sparsity strengths in our data.

In order to address this issue, cross-validation techniques assess significance and reliability based on fitting the PLS models to unseen data. If the model has been overfitted, the latent variable covariances in the unseen data will be low. For example, sPLS using 10-fold cross-validation has been used to successfully model relationships

between 100,000s genetic variables and neuroimaging data (Le Floch et al., 2012). Methods that make use of several folds of the data to calculate better sPLS parameter estimates, such as the hyper-parameter optimisation method proposed by Monteiro et al. hold particular promise for neuroimaging studies in which the sample size is typically low and noise is high (Monteiro et al., 2016). For example, they used this method to identify relationships between multiple items of the Mini-Mental State Examination and grey matter in over 160,000 voxels. However, these methods still require a large sample size. The small training and testing datasets in our dataset of 80 participants resulted in unstable parameter estimates.

Given these challenges, we do not believe that our results reliably indicate the absence of significant relationships between the environment and functional connectivity of the brain at rest. Given the wealth of research in which resting-state connectivity has been found to relate to multiple developmental outcomes and, although limited research, to environmental measures, we would expect there to be associations between the functional connectome and other datasets. For example, Smith et al. found a significant association between multiple aspects of the environment and the resting-state fMRI connectome in an adult sample using very similar methods (Smith et al., 2015). In addition, whilst most of the resting-state literature focuses on fMRI, it has been demonstrated that there is a good consensus between fMRI and MEG resting-state networks (Brookes et al., 2011). In addition, a significant association was found between the environment and the structural connectome in our sample, which has been demonstrated to be closely related to the functional connectome (Hagmann et al., 2008; Hermundstad et al., 2013; Honey et al., 2009). It is also interesting to note that the environmental domains selected as reliable under bootstrap are in line with those identified in the structural connectome chapter, namely the caregiver education and occupation, equivalised income, child health and inverse technology use and discipline. This suggests that it is possible that the PLS model has identified a true association, but that it is not significant due to the issues with overfitting.

Future research will benefit from collecting a larger number of participants. This will enable the application of cross validation techniques in order to address the overfitting apparent in PLS modelling when p>>n. It would also be beneficial to collect longer MEG resting state scans. This would ensure that a greater amount of 'good' data is available after removing the bad sections and would reduce the noise in the estimated partial correlations. In addition, whilst partial correlation has been found to be better at

defining true network structure in comparison to full correlation, it is more difficult to accurately estimate, resulting in higher noise levels. It would therefore be beneficial to also consider full correlations for estimating functional connectomes. It is also possible that the parcellation used might have contributed to the noise in the data. In particular, the parcellation had more parcels than the rank of the MEG data after being processed using the Maxfilter (rank 64) which may have resulted in some parcel time courses being dominated by noise. It would be advisable for future studies to select a parcellation with 64 or fewer parcels. In order to check that the functional connectivity appears to be a sensible estimate, it is also worth using the data to reproduce a seed-based correlation map that has been relatively well established, such as using the primary motor cortex as a seed, to check that the data produces a similar pattern to previous research. Finally, sparse PLS is a new method. To our knowledge this is its first application with multiple imputation, and more work is needed to refine its application and investigate its uses with neuroimaging data.

The importance of the functional connectome in child development holds many exciting research opportunities and further research investigating the relationships between the functional connectome and a child's environment and developmental outcomes is key. Resting-state MEG holds particular potential as it enables us to directly measure electrophysiological activity at good temporal and spatial resolution. This not only allows us to probe the mechanisms by which the environment may affect brain development, and in turn child outcomes but also furthers our understanding of child development in general. Furthermore, studies in which training gains are associated with changes in the functional connectome illustrate the potential to use functional neuroimaging to investigate the effectiveness of childhood interventions for children at risk due to their environment (Barnes et al., 2016). Finally, it has also been demonstrated that it is possible to group children based on aspects of whole-brain organisation such as their connectome (Astle et al., 2018; Bathelt et al., 2018c). Data-driven groupings such as this are an area of particular interest as they allow us to explore why children differ, such as in their response to the environment. It is possible that, one day, this will enable us to better understand why some children are at risk from poor child development whilst others show remarkable resilience to growing up in a deprived environment.

7 Discussion

Growing up in a deprived environment has a profound effect on a child's development. Evidence of this can be seen at multiple levels including key child outcomes such as academic ability and behaviour, cognitive skills and the structural and functional architecture of a child's brain. Identifying the environmental factors that are most strongly linked with these aspects of child development is fundamental for the development of interventions for children at risk. However, investigating the complex relationships between a child's environment and multiple aspects of their development presents a considerable methodological challenge, primarily due to high dimensional, multicollinear and multi-table datasets. As a result, current research has largely focussed on univariate relationships between a limited set of environmental factors and aspects of child development. Identifying the specific domains of a child's environment that are most strongly associated with child development presents a key opportunity that will enable us to develop more detailed theoretical models of the mechanisms and pathways underlying a child's development (Ursache and Noble, 2016). In time, these new models can be tested in fully longitudinal data, as researchers move towards developing and trialling interventions.

A number of methodological advances offer particular promise in addressing these challenges. This study builds on previous research by applying a multivariate approach. Namely, PLS methods enabled us to characterize the environmental profiles most strongly related to multiple aspects of child development. In addition, multi-block PLS allowed us to begin to investigate the possible mechanisms underlying differences in child outcomes associated with their environment. Finally, LASSO feature selection was investigated as a complementary technique to select a smaller set of environmental factors within this profile that are the most useful for predicting child development outcomes. Using these techniques, we employed a multivariate approach to address three key research questions:

1. Which environmental factors most strongly relate to a child's academic ability, behaviour, cognitive ability and neural development?

2. Does the wider environment mediate the relationship between standard measures of SES and child development?

3. How might the environment impact academic and behaviour outcomes? In particular, is this relationship mediated by a child's cognition or the structural and functional connectivity of their brain?

Using a cross-sectional sample of 7-11 children (N=97), we demonstrated the significant importance of applying a multivariate framework to investigate the complex associations between the environment and child development. The results are briefly summarized in the following sections in relation to each of the research questions. These results extend the current research by providing important detail that helps us to understand better why and how the environment is related to child development. This lays the foundation for further longitudinal research and ultimately, we believe that this multivariate approach will enable practitioners and policymakers to better support children at risk from disadvantaged environments.

7.1 Which environmental factors most strongly relate to a child's development?

We moved beyond a simplistic view of a child's environment as measured by standard SES indicators such as caregiver income, education and occupation by considering how multiple environmental domains are associated with multiple aspects of child development. PLS enabled us to summarize these complex associations by identifying pairs of latent variables that covary maximally, each made up of a linear combination of the variables in a given dataset. The environment was significantly related to each aspect of child development. Note that this is with the exception of the difficulties encountered with the functional connectome. There were moderate correlation strengths between the environment and child development latent variables in each pair-wise PLS model, given in Figure 66. Whilst PLS models are correlational and we cannot assume causations, this illustrates the profound and far reaching correspondence between a child's environment and their holistic development, in line with previous research (e.g. (Bradley and Corwyn, 2002b; Brooks-Gunn and Duncan, 1997a; Farah, 2017; McLoyd, 1998; Ursache and Noble, 2016)).
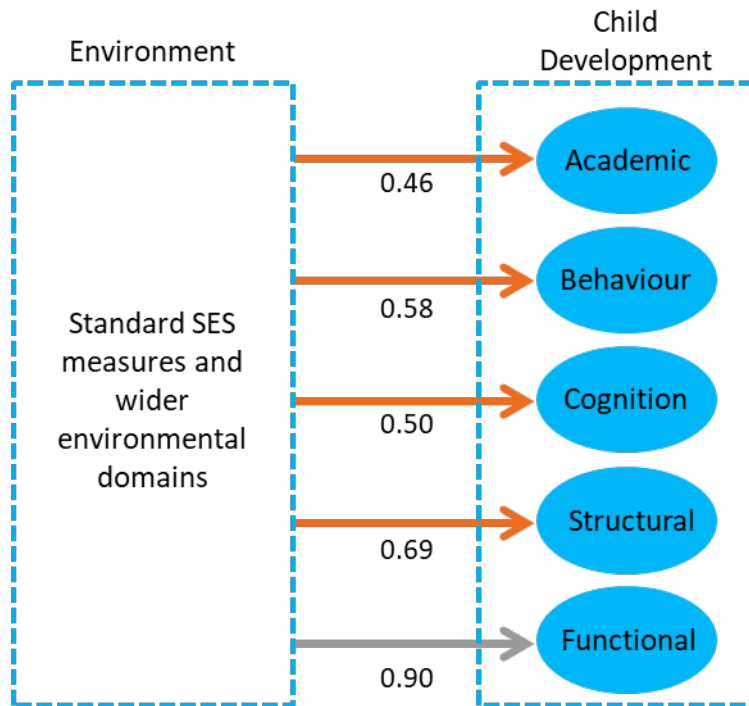
Figure 66 Latent variable correlation between the environment and each aspect of child development. Note that the association between the environment and the functional connectome is not significant under permutation.

The outer weights of the PLS model are proportional to the covariance between a given variable and the latent variable that summarizes the other dataset. Using this property, we investigated the environmental factors most strongly related to a child's development. Figure 67 summarizes the results across each level of child development. The colour represents the covariance between each environmental domain and each aspect of child development. The stars denote domains selected as reliably related to child development in the PLS model.
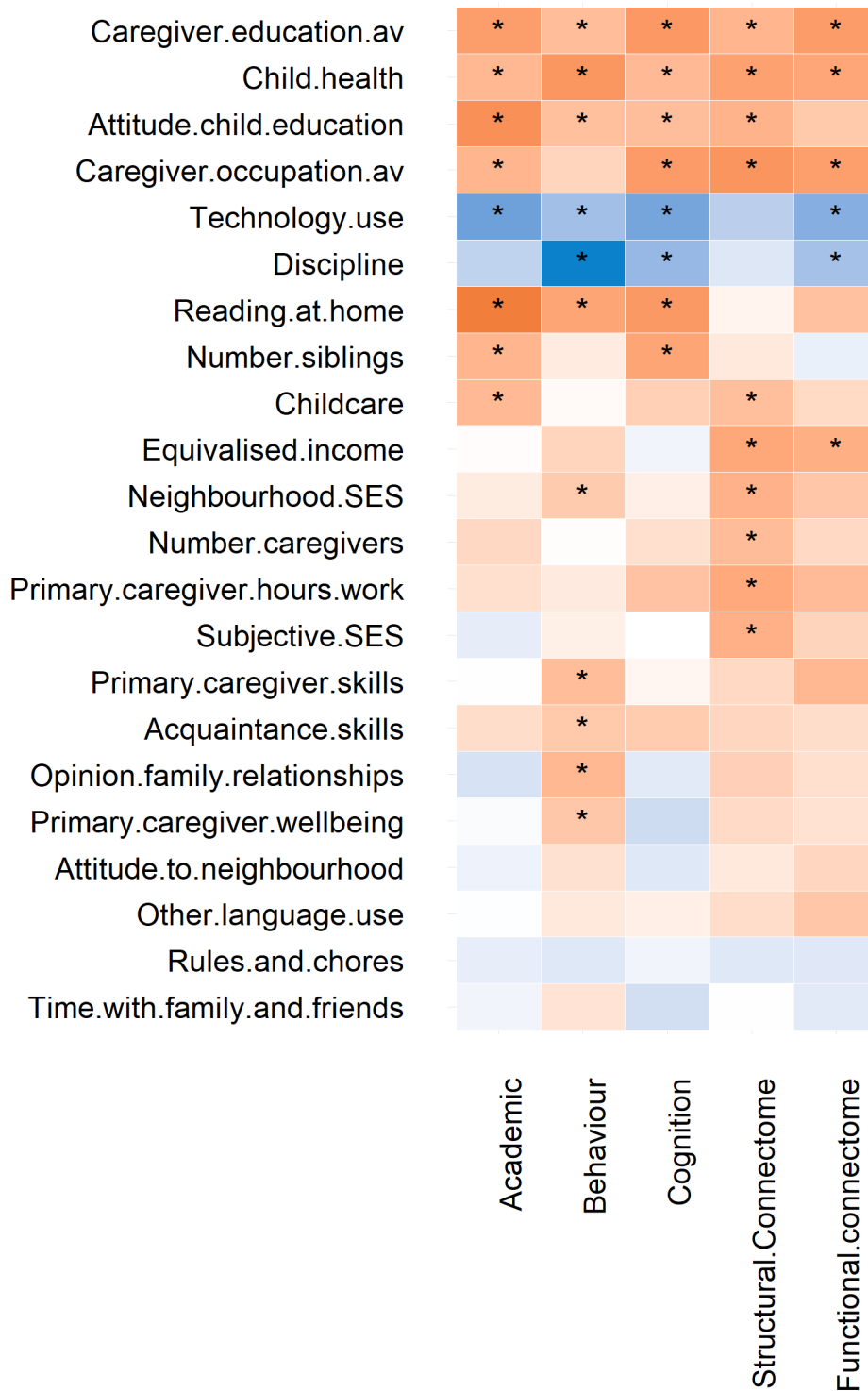
Figure 67 Outer weights for each environmental domain in relation to each aspect of child development , proportional to the covariance between the domain and child development latent variables. Environmental domains found to reliably covary with the child development measures are denoted by a star. The domains are ordered according to the number of times they are selected as reliable across the different child development measures.

A number of themes emerge from examining the bigger picture associations between the environment and child development. Perhaps the most surprising result is that several aspects of a child's wider environment were found to be similarly or, in some cases, more strongly associated with child development than standard measures of SES. Indeed, whilst caregiver education and occupation were associated with all aspects of child development, equivalised income was not found to associate with academic attainment, behaviour or cognition. In other words, several participants demonstrated an incredible resilience to growing up in relative poverty, and these individual differences are likely to be driven by wider aspects of a child's environment. Additionally, none of the SES measures were selected as being the most predictive of any of the child outcomes in the supplementary LASSO analysis, indicating that the covariance between SES and child outcomes was captured by differences in these wider environmental domains.

For example, better child health was identified as one of the most strongly associated environmental domains across each level of child development. Encompassing several of the physical aspects of a child's environment such as their nutrition, sleep or outdoor play, the importance of child health is in line with evidence across a wide body of literature (for example, see Aarons et al., 2008; Engle et al., 2007; Gregg et al., 2007; Sibley and Etnier, 2003; Sorhaindo et al., 2006; Wilson et al., 2012). Two domains indexing educational attitudes and behaviours, attitude to education and reading at home, were also found to be particularly important at multiple levels of child development. Interestingly, these educational attitudes are consistently found to be some of the most powerful predictors of child development (Bradley et al., 2001; Education standards research team, 2012; Gregg et al., 2007; Loeber and Stouthamer-Loeber, 1986; OECD, 2002; The Reading Agency, 2015). In addition, some domains were inversely related to the child development measures, namely technology use and discipline. Primarily, technology use indicates a child's access and use of technology for entertainment and the discipline domain largely measures harsher discipline techniques. These negative associations have also been identified in literature (e.g. see Bavelier et al., 2010; Harris et al., 2017; Hastings et al., 2009; Mistry et al., 2007; Schmidt and Vandewater, 2008 for technology use and Bailey et al., 2009; Byford et al., 2012; Gregg et al., 2007; Wilson et al., 2012 for discipline). Notably, each of these domains were also selected as being the most predictive of the child outcomes in the supplementary LASSO analysis.

Whilst SES measures typically provide useful proxy measures of a child's environment, these results highlight the importance of building on these simplistic measures by investigating individual differences in a child's environment in more detail. Treating children from one income bracket as a homogeneous group whilst failing to consider the wider environment may, in fact, explain some of the inconsistencies in the SES literature. For example, for almost every significant association found, a non-significant effect has also been reported between standard SES measures and child developmental markers in other studies. For example, several papers have reported finding non-significant associations between SES and global white and grey matter volume (Butterworth et al., 2012b; Jednorog et al., 2012; Lange et al., 2010), hippocampus (Butterworth et al., 2012a) and amygdala volume (Butterworth et al., 2012a; Hanson et al., 2011a; Noble et al., 2012b, 2015), white matter integrity (Jednorog et al., 2012), resting state activity (Brito et al., 2016; Tomalski et al., 2013) and cognitive skills (Butterworth et al., 2012b; Czernochowski et al., 2008b; Jednorog et al., 2012; Kishiyama et al., 2009, 2009; Lupie et al., 2001; Noble et al., 2005). Other studies find an association between differing aspects of SES. For example, studies have found that income and not parent education is related to hippocampus volume (Hanson et al., 2011b; Luby et al., 2013; Noble et al., 2012c) whereas others have found that parent education and not income is associated (Noble et al., 2015). The mixed findings are likely to be due to unmeasured differences in a child's wider environment. Investigating a wider array of environmental measures helps us to understand why some samples appear to be resilient to low SES environments and ensures that we do not miss key environmental effects.

In addition, the complex set of associations illustrates the importance of using multivariate methods to consider a wide array of environmental factors simultaneously. For example, we might identify a significant pairwise association between a given aspect of the environment and child development. However, measuring limited environmental markers, an approach common in the literature, masks the fact that a child's environment is governed by a variety of factors that interact and covary. A significant association may instead be driven by strong correlations between the environmental measure, the child development measure and a third unmeasured environmental variable. For example, technology use was found to be reliably inversely related to cognitive skills. However, technology use also covaries with several other strong predictors of cognitive skill and controlling for these has been shown to remove

the association between technology use and cognition (Schmidt et al., 2009; Schmidt and Vandewater, 2008). In short, a strong possibility is that our particular focus on technology for entertainment means that the factor acts as a proxy measure for a wider environmental profile, not a direct pairwise relationship in its own right. Thus, a great deal of care must be applied when interpreting correlations between aspects of the environment and child development. By considering a wide range of related environmental variables using multivariate methods, we have captured this complexity. We identified a set of candidate environmental domains that have the potential to be causally linked to child development, developing an evidence base for further investigation.

Furthermore, these results also highlight the importance of considering multiple levels of child development. This allows the researcher to identify profiles associated with specific outcomes, and those elements of a profile that appear to have more widespread relationships across multiple outcome measures. For example, it is particularly interesting that some of the environmental markers such as child health, attitude to education and inverse technology are associated across the multiple levels of development. These present key potential targets for interventions as their impact is likely to be far reaching across child development markers. These can be contrasted from relationships that are more specific. For example, as shown in Figure 61, social resources such as caregiver and acquaintance skills, subjective opinion of family relationships and caregiver wellbeing are only reliably related to child behaviour. Identifying the domains associated with specific aspects of development pinpoints potential environment factors of risk and resilience that could be future targets for interventions designed for particular childhood outcomes.

Finally, different levels of child development may be differentially sensitive to environmental effects. For example, a set of environmental domains appear to be more important for the brain architecture measures than the child outcomes or cognitive skills. Specifically, many of these domains are closely related to SES such as equivalised income, subjective and neighbourhood SES and caregiver work hours. It is possible that differences in certain markers of child development may precede other markers. For example, neural measures might be more sensitive to early changes in a child's environment, which might take longer to emerge as measurable differences in child outcomes. Not only would these differences be missed if we did not consider

'lower' levels of development, but these may provide early biological indicators of environmental risk or resilience.

In summary, these results build on the current SES literature by illustrating the importance of using multivariate approaches when investigating the relationship between the environment and child development. A number of environmental domains were found to reliably relate to several aspects of a child's development. The relationships identified provide important detail that enables us to better develop theories about which aspects of a child's environment are associated with different levels of child development. Notably, several of these environmental domains have the potential to be easier to address than standard markers of SES, such as health, attitudes to education and parenting behaviours, making them promising potential targets for intervention studies.

7.2 Does the wider environment mediate the relationship between standard measures of SES and child development?

Whilst the first research question addresses how the environment is related to child development, our second aim was concerned with why a child's SES is associated with their development. Given the importance of aspects of a child's wider environment outlined above, it is likely that SES has an impact on child development by driving differences in the wider environment. For example, higher caregiver education is likely to result in better attitudes to education and increased reading at home which in turn might impact the childhood outcomes. We investigated whether the association between standard measures of SES (income, occupation and education) and each aspect of child development is mediated by the other environmental domains, using multi-block PLS. By applying recent methodological advances in the analysis of multiple large multivariate datasets, we were able to apply path modelling without the typical constraint that the number of samples must be larger than the number of variables. The resulting path coefficients and related latent variable outer weights identify candidate mechanisms underlying why SES is associated with the different aspects of child development. This is particularly important as it is these mediators that are often targeted by intervention studies. Consequently, providing a better understanding of these relationships will help to develop more effective and efficient interventions.

The wider environment significantly mediated the relationship between SES and each aspect of child development (see Figure 68). In fact, the wider environmental domains

fully explained the relationships between SES and academic attainment, behaviour and cognitive skills. This suggests that SES related differences in key childhood outcomes and their underlying cognition are due to SES driven differences in a subset of the environmental domains measured. Interestingly, the relationship between SES and the structural and functional connectome was still significant after the wider environment had been taken into account. This means that the SES measures explained additional variance in the architecture of the brain not captured by the environmental domains. This suggests that there may be other mechanisms by which SES might affect the brain that are not captured in the domains measured.
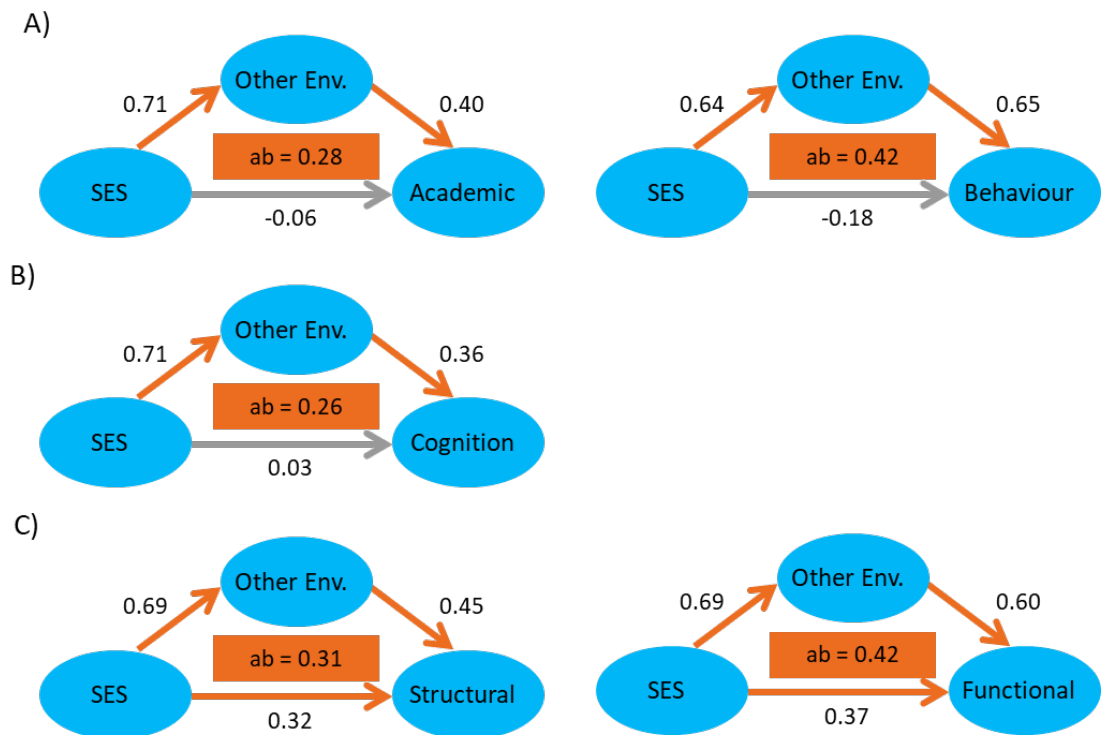
Figure 68 Path models investigating the mediating effect of the wider environment on the relationship between SES and child outcomes (A), cognition (B) and their structural and functional brain connectome (C).

The outer weights of the PLS model are proportional to the covariance between a given variable and the latent variables that summarize the other datasets. For example, the outer weight for, say, technology use is proportional to the total covariance between technology use and both the SES and child development latent variable. This makes interpretations more challenging as a high outer weight does not necessarily mean that that a variable covaries with both of the other datasets. This is illustrated by the fact that more variables are identified that load reliably onto the mediating latent variable than

for the two-block PLS models, summarized in Figure 69. Indeed, some variables that are not found to relate to child development are identified as reliable because they have a strong relationship to SES such as use of other languages or time with family and friends.

However, using the Pearson correlation tables from sections 3.3.1 and 4.3.1 and the PLS models from the previous section allows us to see that the domains with the highest outer weights are those that covary with both the standard SES measures and the aspects of child development. In particular, these include those selected as most reliable in the two-block PLS models; child health, attitude to education, reading at home and inverse technology use and discipline. This suggests that these domains are candidate mechanisms by which SES might impact a child's development. This provides further evidence that these domains might be particularly effective targets for interventions aimed at supporting children at risk due to growing up in a deprived environment.

Figure 69 Outer weights for the wider environment latent variable in the path analysis. Environmental domains found to reliably load onto the latent variable are denoted by a star. The domains are ordered according to Figure 67

7.3 How might the environment impact academic and behaviour outcomes?

The different levels of child development measured are not independent. For example, a child's cognitive abilities and, at a lower level, the neural architecture of their brain provide the foundation for the development of many skills including academic ability (Deary et al., 2007; Oakhill and Cain, 2012; Pugh et al., 2001; Tau and Peterson, 2010; Welsh et al., 2010) and behaviour (Helton et al., 2018; Muetzel et al., 2018; Pu et al., 2017; Willoughby et al., 2011). Furthermore, a change in one aspect of development is likely to cascade across the other aspects of development over time. Given the significant relationships between the environment and each aspect of child development, we hypothesized that the environment might impact childhood outcomes via changes in a child's cognitive skills or brain architecture. However, to date, little has been done to bridge the different levels of child development in order to investigate the candidate mechanistic pathways by which the environment might exert an effect. We used multi-block PLS to explore whether cognition or brain architecture mediated the relationship between the environment domains and the key child outcomes.

The path models for each analysis are summarised in Figure 70. We found that both cognitive skills and the structural connectome fully mediated the relationship between the environment and a child's academic ability. Note that the functional connectome also mediated the relationship but given that some of the path coefficients are greater than one due to the high dimensionality and multicollinearity of the dataset, further research is needed. This suggests that a child's cognition and neural architecture provide mechanistic pathways underlying the attainment gap in academic ability associated with a child's environment. In line with the limited research available (Alves et al., 2017; Crook and Evans, 2014; Dilworth-Bart, 2012; Gullick et al., 2016; Hair et al., 2015; Nesbitt et al., 2013; NICHD Early Child Care Research Network, 2003), this provides key evidence towards developing a theoretical understanding as to how the environment impacts a child's development. In particular, this is crucial for developing interventions to support children most at risk of poor academic development due to growing up in a deprived environment. For example, it is possible that providing better support for academic skills may be further enhanced by supporting neural and cognitive development earlier in life. It may, for example, be beneficial to provide training in a cognitive skill such as working memory, which in turn will better support learning. Interventions like this have so far shown that functional and structural connectivity can

be enhanced, and these domains of cognition boosted, but evidence for far transfer remains weak (Astle et al., 2015; Bergman Nutley and Söderqvist, 2017; Cao et al., 2016; Sala and Gobet, 2017; Takeuchi et al., 2010). As we further develop our knowledge about which specific elements of cognition and brain organisation are most strongly linked with resilience over developmental time, these may become the targets for future interventions.

We were surprised to find that cognition and the structural connectome did not mediate the relationship between the environment and behaviour. Note that this was the first study, to our knowledge, investigating these mechanistic pathways. This suggests that the environment might be associated with child behaviour via different mechanisms to the impact on academic ability. One possible reason for this is that behavioural differences might be driven by differences in environmental domains that are not associated with the cognition or brain structure. For example, a number of domains largely concerning social resources were primarily associated with behaviour and not the other aspects of child development. This is particularly evident for the cognition pathway in which the environment explained significant variance in child behaviour even when cognition was held constant. This highlights the potential importance of social resources for child behaviour. It also illustrates the need for caution when interpreting correlational relationships. For example, the well documented associations between cognition and behaviour (Flouri et al., 2015; Harpur et al., 2015; Helton et al., 2018; Hinshaw, 1992; Masten et al., 1999; Menting et al., 2011; Willoughby et al., 2011) might in fact be due to underlying correlation with the environment rather than a direct relationship. However, it is also possible that the cognitive and brain measures used might not be as sensitive to behavioural differences as other possible measures. For example, it is likely that 'hot' executive functions, which have been found to be closely linked to both behavioural outcomes and child SES might instead provide a mechanistic pathway underlying the relationship between the environment and child behaviour. In addition, it is possible that behavioural difficulties arise from localized differences in a child's brain which might not be visible when considering the whole-brain connectome. Finally, behaviour was rated subjectively through parent questionnaires. Whilst these are standard questionnaires used widely, it is possible that these questionnaires do not measure the aspects of behaviour that might be more dependent on the underlying neural circuitry and cognitive ability and a significant mediation may be found using other, more rigorous assessments of behaviour, such as

rater-blind observations. As a result, we cannot rule out that there are cognitive or neural pathways that mediate the relationship between the environment and a child's behaviour. But our results do show that the mediating pathways for academic attainment and behaviour are different from one another.
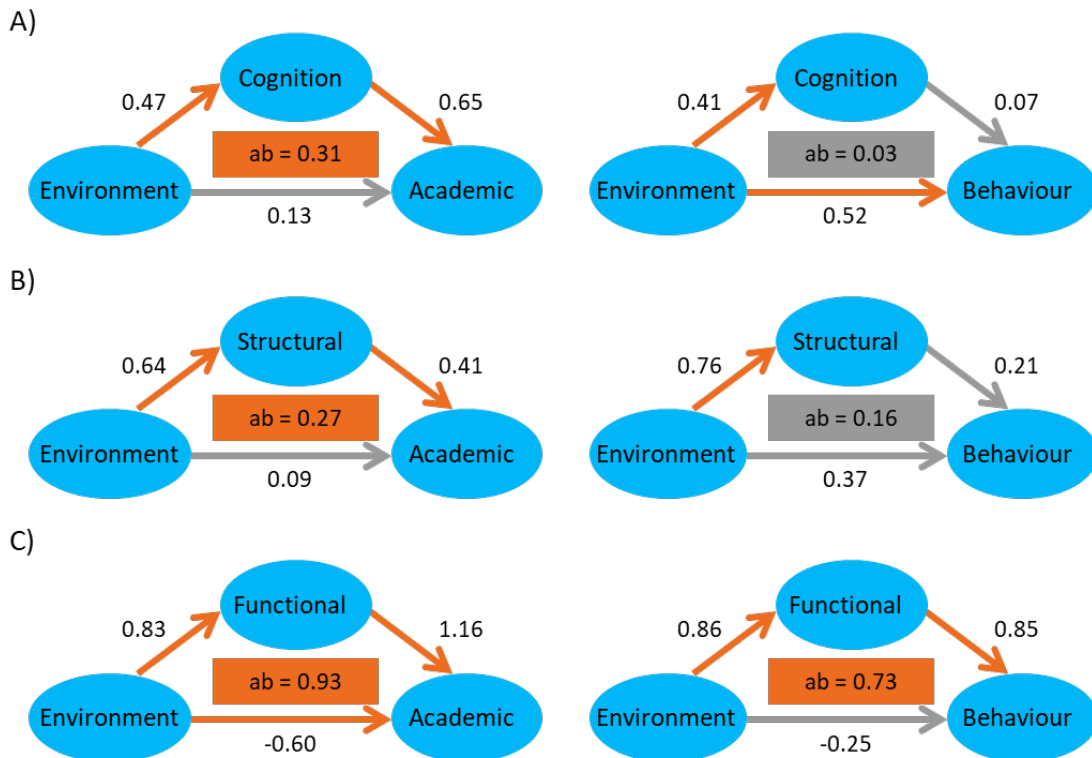
Figure 70 Path models investigating the mediating effect of a child's cognitive ability (A), structural connectome (B) and functional connectome (C) on the relationship between the environment and child academic and behaviour outcomes

The outer weights for the environmental domains are summarized as before in Figure 71. Note that only the first two columns relate to PLS models in which a significant mediation effect was found. However, the outer weights are the combined covariance between each domain and both the child outcomes and the mediatory aspect of child. Therefore, those with the highest weights are likely to relate to the multiple aspect of child development and so these give us an indication of potential environmental domains that are most likely to relate to child outcomes via the cognitive skills or brain architecture. As before, we can see that there is the same key subset of domains that have high covariance with the other datasets including child health, attitudes to education and reading at home and inverse technology use and child discipline. This suggests that these domains have the potential to impact childhood outcomes via differences in a child's cognition or neural architecture.
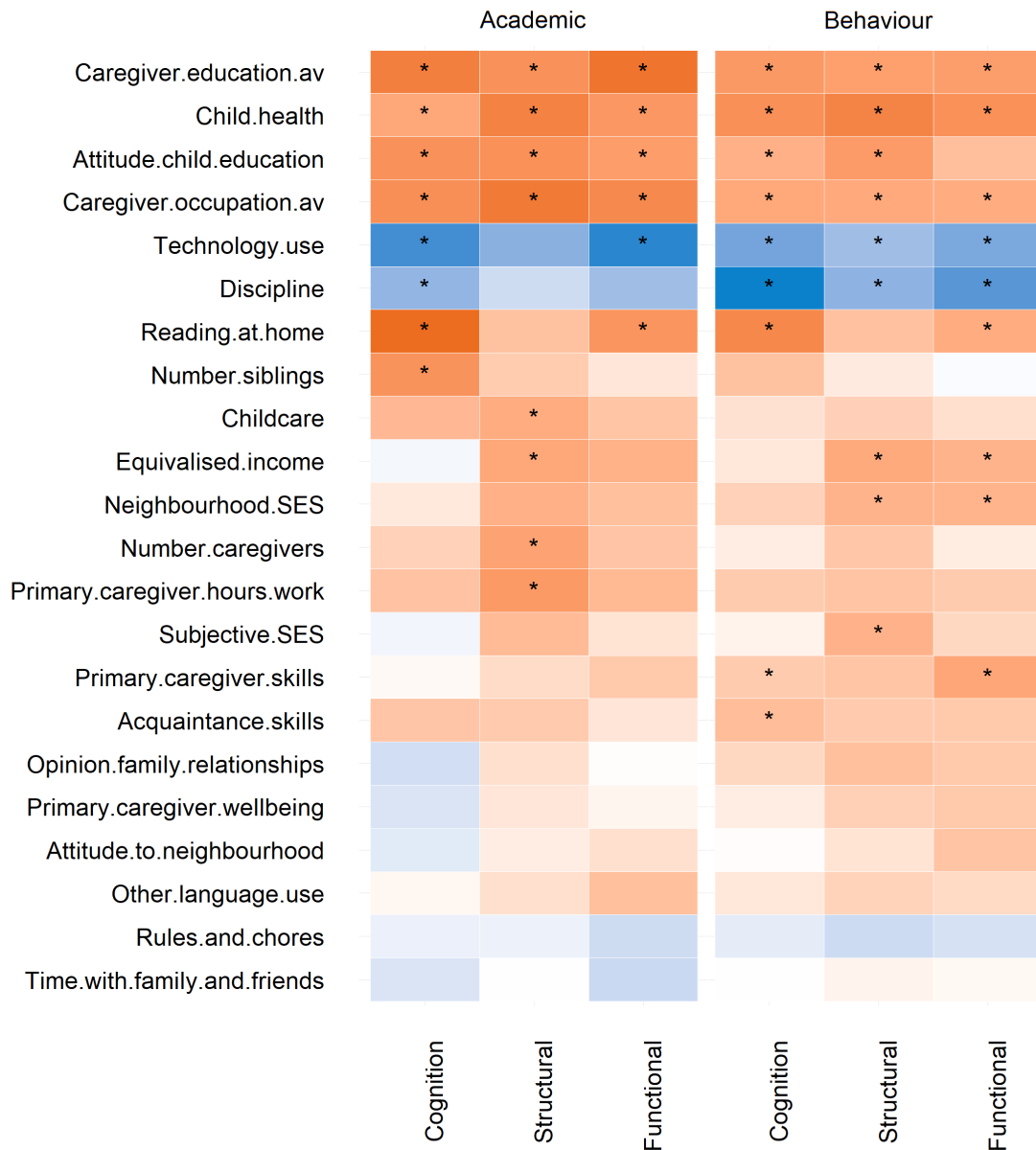
Figure 71 Outer weights for the environment domains in the path analysis investigating whether cognitive skills or the structural and functional connectome reliably mediate childhood academic and behaviour outcomes. Domains found to reliably load onto the latent variable are denoted by a star and are ordered according to Figure 67

7.4 Limitations

There are number of limitations in this study that could be addressed in future work. Firstly, more could be done to improve the environmental domains. There is very limited research investigating the wider aspects of a child's environment might be related to their development, and so there was relatively little theoretical evidence to guide our choice of domains. As a result, we opted to include a wide variety of domains

to ensure that we capture any associations. Where possible, we used standard and widely used questionnaires to measure these aspects, such as the resource generator for caregiver and acquaintance skills (Webber and Huxley, 2007), the MacArthur Scale of Subjective Social Status (Adler et al., 2000) and the English Indices of Deprivation for neighbourhood SES (Office for National Statistics, 2015). However, due to the lack of available questionnaires regarding a child's wider environment, we had to create our own in-house set of questionnaires. Not only does this make it difficult to compare our results to other studies or to replicate this study, but they have not been validated or refined using a large sample representative of the population. In addition, investigating such a wide number of environmental domains limited the number of items that could be measured for each domain due to time constraints. As a result, some of the domains may not be reliably measured. For example, once we removed all of the childcare questions that had highly skewed answers, only two questions remained to use in the factor analysis for the childcare domain. It is possible that the lack of association between domains such as rules and chores or primary caregiver wellbeing might be because the items used do not capture these underlying factors well. Therefore, further research will benefit from employing standard questionnaires such as the HOME inventory (Bradley and Caldwell, 1977; Elardo and Bradley, 1981; Totsika and Sylva, 2004) or the construction and validation of a new questionnaire that taps multiple aspects of a child's environment.

Secondly, we used a factor analysis to combine questionnaire items to create an underlying factor score. However, our quasi-data-driven approach required several a priori assumptions – we had to assume that the child's environments are indeed made up of the underlying constructs chosen and that the items load onto the particular domains selected. One alternative would be to take a purely data driven approach by inputting all items into one factor analysis such as the method applied by Markus Jenkins et al. (Marcus Jenkins et al., 2013) but the resulting factors are likely to be more difficult to interpret. In addition, we chose to keep only the first factor for each domain, as this also improves interpretability. However, this means that some of the variance captured by the question items is lost. Given this, there is a possibility that the variance captured is not the variance most strongly associated with the environmental domains. Another possibility could be to include each of the questionnaire items as variables in the PLS models, in a similar way to the method used in the functional connectome study by Smith et al. (Smith et al., 2015). Note that this would involve hundreds of question

items and would be more difficult to interpret. In addition, combining multiple items in a factor analysis is likely to provide a more reliable measure of an underlying environmental domain than a single questionnaire item. However, further research is needed to develop theoretical models of underlying environmental constructs to better inform item selection and data reduction methods designed to measure these domains.

Thirdly, one of the biggest constraints is that we cannot assume causality in any of the relationships found. Thus, we cannot say that differences in the environment cause the differences in the aspects of child development identified in this study. The environmental domains selected as reliable covariates of the child development measures are typically moderately correlated with each other. Whilst it is possible, if not probable, that some of these domains causally impact child development, we cannot rule out the possibility that they may only be associated due to a third covarying environmental domain. For example, technology use might be strongly inversely associated with child outcomes simply because it is a particularly good proxy measure of other aspects of a child's environment. Technology use is particularly strongly correlated with many of the other domains frequently identified as associated with child development. Indeed, this study highlights the need for caution by identifying so many domains that do covary with the child outcomes and highlights the need for powerful multivariate techniques capable of incorporating multiple covarying environmental measures. In addition, correlations may also be driven by factors that are not measured in this study. For example, one of the most famous debates surrounding this area of research is the question of whether differences in a child's environment are driven by nature vs nurture. It is possible that genetic factors are causally related to both the environment domains and the childhood outcomes, resulting in the associations seen.

There are similar limitations specific to the mediation analyses. Namely, path analyses require assumptions about the direction of the paths between the datasets. This is particularly an issue for cross-sectional studies such as there is no way to test these assumptions. For example, we assumed that the environment exerts an impact on cognition which in turn impacts child behaviour and so investigated whether cognition mediated this relationship. However, the path coefficients between the environment and cognition and between the environment and behaviour were both identified as reliable under bootstrap, but the path between cognition and behaviour was not reliable. Instead, it is possible that cognition might impact the environment resulting in changes in child behaviour. When we reversed the arrows, we found that the environment mediated the

relationship between cognition and behaviour. This was particularly challenging in this analysis as a few (but not all) of the domains have the potential to be 'caused' by cognitive skills such as the amount a child chooses to read at home. It is in fact probable that some arrows go both ways; i.e. an increase in one aspect of child development due to the environment is likely to facilitate improvements in the environmental domains. Such positive feedback loops might also exist between the child development measures. Further studies are needed to develop stronger theories about how different aspects of the environment and child development relate and interact. For example, longitudinal datasets that combine these rich measures with multiple time points would provide an ideal dataset.

The outer weights from the first component of a PLS model are particularly useful in identifying variables that are key covariates of another dataset from amongst a number of other variables. However, another limitation related to PLS is that the higher order components become very difficult to interpret. This is because the variance explained in the first component is typically removed from the datasets (deflation) before fitting the PLS model to get the second component and so on. For example, a second component outer weight for an environmental domain is proportional to the variance in the child development measure after the primary covariance identified by the first factor has been removed. It is possible to use other deflation techniques. For example, it is common to only deflate one of the datasets, leaving the other intact in two-block PLS. However, this becomes more difficult with multi-table datasets. Alternatively, it has recently been suggested that the latent variables could be rotated to create a simpler factor structure, using methods similar to those used in factor analyses. It would be interesting to see whether the environmental domains group into subsets that have a specific association with the measures of child development.

Furthermore, it is also more difficult to interpret outer weights in the mediation analyses as each other weight is proportional to the sum covariance between that variable and the latent variables for both of the other datasets. This means that an outer weight can be entirely driven by a strong association with one dataset. Therefore, high outer weights in the mediating dataset do not necessarily mean that these variables facilitate the mediation effect. We addressed this by using the two-block PLS results and Pearson correlations to guide the interpretation. However, our results have pinpointed potential candidate variables for the mediation effects and it would be useful to consider these domains in isolation to identify whether they mediate the relationships investigated.

Finally, the last results chapter highlights a key challenge that is becoming increasingly important to address as our data sets get bigger. That is, statistical methods such as PLS are unsuitable when the number of variables far exceeds the number of participants. For example, the resting state MEG functional connectome consisted of a very high number of variables that were also relatively noisy. As a result, the PLS models were not significant under permutation but had very small bootstrap errors resulting in the potential for both false negatives and false positives. Introducing sparsity as part of the PLS modelling presents a particularly promising solution. However, these methods are relatively new and remarkably few papers have applied these techniques. It is imperative that more research is conducted to develop and verify these methods, particularly in the context of neuroimaging data which typically suffers from smaller sample sizes.

7.5 Future directions

The research available to date has primarily focussed on pairwise relationships between limited aspects of a child's environment and their development. By applying a multivariate approach, we have expanded this by identifying multiple aspects of a child's environment that are reliably related to child development at multiple levels. In short, we have attempted to identify environmental profiles most strongly linked with particular key developmental outcomes. We have also explored the potential mechanistic pathways underlying these relationships. These results provide a valuable and detailed evidence base for future studies to draw on. Furthermore, we have demonstrated the far-reaching impact of a child's environment across multiple levels of their development, highlighting the importance of finding better ways to support children most at risk from growing up in a deprived environment. Going forward, a priority is understanding whether these associations and mediation effects are causal or simply correlational.

As a next step, it would be particularly useful to expand on this cross-sectional analysis by repeating the study in the same participants over time. Such longitudinal analyses allow us to investigate whether individual differences measured at the first time point are predictive of the rate of change of other variables over time. For example, we might investigate whether a child's wider environment at the first time point predict the rate of development of the child's academic and behaviour outcomes, cognition or neural architecture. If a variable was a good predictor of later developmental measures after

taking into account the current stage of development, this would provide stronger evidence that that variable exerts a causal effect on child developmental.

By developing our theoretical understanding of how and why the environment is associated to child development, we will be better able to identify key targets for further intervention studies. These are the gold standard for establishing causality. For example, we might investigate how one aspect of a child's environment impacts child development downstream. Ultimately, this will enable practitioners and policy makers to identify the most effective interventions to support children at risk of poor development due to their environment.

In addition, we have highlighted the importance of neuroimaging studies for developing our theoretical understanding of how and why the environment is associated with a wide variety of developmental outcomes. By investigating the underlying neural architecture of a child's brain, we gain insight into the potential mechanistic pathways by which the environment might impact child development at multiple levels. In particular, we have identified the need for studies to move beyond global averages or localised differences towards considering individual connections across the whole-brain connectome. We found that associations between the connectome data and the environment and child outcomes were widespread across the brain rather than localised to specific brain regions. However, as highlighted in the analysis of the functional connectome, PLS methods are prone to overfitting when the number of variables far exceeds the number of samples. Going forward, it will be important to collect data from larger samples to enable the use of methods such as cross-validation techniques to ensure the PLS models generalize to other samples. In relation to this, collecting quality data is also crucial. For example, we found it challenging to collect a longer resting state MEG scan with children. However, increasing the number of time points or reducing the distance between the MEG sensors and the scalp (Boto et al., 2017) would improve the signal to noise ratio and hence provide a better estimation of the functional connectome.

Furthermore, we have outlined methodological techniques that can be applied to multiple imaging modalities. For example, it would be interesting to investigate the structural brain architecture further by using other tractography measures or connectomes based on the covariance between the structural properties of different brain regions such as the surface area or thickness. Resting state connectivity is most widely measured using fMRI. Whilst MEG imaging offers several advantages, fMRI has a higher spatial resolution and can reliably measure functional connectivity in

subcortical brain regions. Investigating the resting state functional connectome using this would provide a key complementary analysis to the MEG connectome results. In addition, it would also be useful to investigate the functional connectome whilst participants complete specific tasks, such as those that tap into academic ability, behaviour and cognitive skills, as these may emphasize individual differences in the connectome that are specific to key aspects of child development. Finally, recent developments in the application of graph theory offer us a wealth of information about the underlying properties of these complex networks (Rubinov and Sporns, 2010). For example, individual differences in measures such as the node degree, clustering coefficient, efficiency, hub architecture or modularity are increasingly identified as key properties of child development (Bathelt et al., 2018d; Hwang et al., 2013; Kaiser, 2017; Sporns, 2013). It will be interesting to consider how these underlying properties of the structural and functional connectomes are associated with the environment and whether they mediate the relationship between the environment and child development. Thus, extending this analysis to other aspects of a child's brain architecture presents a key research opportunity for developing our understanding as to why and how the environment that a child grows up in is so profoundly associated to their child development.

8 CONCLUSIONS

In conclusion, we highlight the importance of applying a multivariate framework to investigate the complex relationships between a child's environment and multiple aspects of their development. To date, the considerable methodological challenges inherent in investigating these complex relationships has resulted in the current research largely focussing on univariate relationships between a limited set of environmental factors and aspects of child development. We illustrate the use of multivariate techniques such as PLS and LASSO feature selection to address these challenges. Multiple environmental domains were found to be reliably related to each aspect of child development. In addition, the wider environmental domains mediated the association between SES measures and child development. Furthermore, cognition and the structural connectivity of a child's brain mediated the association between the environment and academic outcomes, but not behaviour outcomes. These results extend the current research by providing important detail that helps us to understand better why and how the environment is related to child development. This lays the foundation for further longitudinal research and the development of more effective interventions. Ultimately, we believe that this multivariate approach will enable practitioners and policymakers to better support children at risk from disadvantaged environments.

9 REFERENCES

Aarons, G.A., Monn, A.R., Leslie, L.K., Garland, A., Lugo, L., Hough, R.L., Brown, S.A., 2008. The Association of Mental and Physical Health Problems in High-Risk Adolescents: A Longitudinal Study. J. Adolesc. Health Off. Publ. Soc. Adolesc. Med. 43, 260–267. https://doi.org/10.1016/j.jadohealth.2008.01.013

Ackerman, B.P., Brown, E.D., 2006a. Income poverty, poverty co-factors, and the adjustment of children in elementary school. Adv. Child Dev. Behav. 34, 91–129.

Ackerman, B.P., Brown, E.D., 2006b. Income poverty, poverty co-factors, and the adjustment of children in elementary school. Adv. Child Dev. Behav. 34, 91–129.

Adams, L., 2015. Utilization of the Alabama Parenting Questionnaire across Family Structures: Do the Same Constructs Apply? LSU Dr. Diss.

Adler, N.E., Epel, E.S., Castellazzo, G., Ickovics, J.R., 2000. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. Health Psychol. Off. J. Div. Health Psychol. Am. Psychol. Assoc. 19, 586–592.

Alloway, T.P., 2007. Automated working memory assessment: Manual. Pearson.

Alloway, T.P., Alloway, R.G., 2010. Investigating the predictive roles of working memory and IQ in academic attainment. J. Exp. Child Psychol. 106, 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alves, A.F., Gomes, C.M.A., Martins, A., Almeida, L. da S., 2017. Cognitive performance and academic achievement: How do family and school converge? Eur. J. Educ. Psychol. 10, 49–56. https://doi.org/10.1016/j.ejeps.2017.07.001

Andreou, C., Faber, P.L., Leicht, G., Schoettle, D., Polomac, N., Hanganu-Opatz, I.L., Lehmann, D., Mulert, C., 2014. Resting-state connectivity in the prodromal phase of schizophrenia: insights from EEG microstates. Schizophr. Res. 152, 513–520. https://doi.org/10.1016/j.schres.2013.12.008

Andrews, J., Robinson, D., Hutchinson, J., 2017. Closing the Gap?: trends in educational attainment and disadvantage. August 2017. Educ. Policy Inst.

Aneshensel, C.S., Sucoff, C.A., 1996. The neighborhood context of adolescent mental health. J. Health Soc. Behav. 37, 293–310.

Anyaegbu, G., 2010. Using the OECD equivalence scale in taxes and benefits analysis. Economic & Labour Market Review 4.

Astle, D.E., Barnes, J.J., Baker, K., Colclough, G.L., Woolrich, M.W., 2015. Cognitive training enhances intrinsic brain connectivity in childhood. J. Neurosci. Off. J. Soc. Neurosci. 35, 6277–6283. https://doi.org/10.1523/JNEUROSCI.4517-14.2015

Astle, D.E., Bathelt, J., CALM team, Holmes, J., 2018. Remapping the cognitive and neural profiles of children who struggle at school. Dev. Sci. e12747. https://doi.org/10.1111/desc.12747

Australian Institute of Family Studies, 2017. Growing Up in Australia: The Longitudinal Study of Australian Children [WWW Document]. URL http://www.growingupinaustralia.gov.au/ (accessed 6.8.17).

Avants, B.B., Tustison, N.J., Johnson, H.J., 2009. ANTS: Advanced Open-Source Normalization Tools for Neuroanatomy. Penn Image Computing and Science Laboratory. Penn Image Computing and Science Laboratory.

Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J., 2011. Multiple Imputation by Chained Equations: What is it and how does it work? Int. J. Methods Psychiatr. Res. 20, 40–49. https://doi.org/10.1002/mpr.329

Baccalá, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. 84, 463–474. https://doi.org/10.1007/PL00007990

Bailey, J.A., Hill, K.G., Oesterle, S., Hawkins, J.D., 2009. Parenting practices and problem behavior across three generations: Monitoring, harsh discipline, and drug use in the intergenerational transmission of externalizing behavior. Dev. Psychol. 45, 1214–1226. https://doi.org/10.1037/a0016129

Baker, B.L., McIntyre, L.L., Blacher, J., Crnic, K., Edelbrock, C., Low, C., 2003. Pre-school children with and without developmental delay: behaviour problems and parenting stress over time. J. Intellect. Disabil. Res. JIDR 47, 217–230.

Barch, D., Pagliaccio, D., Belden, A., Harms, M.P., Gaffrey, M., Sylvester, C.M., Tillman, R., Luby, J., 2016. Effect of Hippocampal and Amygdala Connectivity on the Relationship Between Preschool Poverty and School-Age Depression. Am. J. Psychiatry 173, 625–634. https://doi.org/10.1176/appi.ajp.2015.15081014

Barch, D.M., 2017. Resting-State Functional Connectivity in the Human Connectome Project: Current Status and Relevance to Understanding Psychopathology. Harv. Rev. Psychiatry 25, 209–217. https://doi.org/10.1097/HRP.0000000000000166

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., Dant, C.C., Reiss, A.L., 2005. White Matter Development During Childhood and Adolescence: A Cross-sectional Diffusion Tensor Imaging Study. Cereb. Cortex 15, 1848–1854. https://doi.org/10.1093/cercor/bhi062

Barnes, J.J., Nobre, A.C., Woolrich, M.W., Baker, K., Astle, D.E., 2016. Training Working Memory in Childhood Enhances Coupling between Frontoparietal Control Network and Task-Related Regions. J. Neurosci. Off. J. Soc. Neurosci. 36, 9001–9011. https://doi.org/10.1523/JNEUROSCI.0101-16.2016

Barnes, J.J., Woolrich, M.W., Baker, K., Colclough, G.L., Astle, D.E., 2015. Electrophysiological measures of resting state functional connectivity and their relationship with working memory capacity in childhood. Dev. Sci. https://doi.org/10.1111/desc.12297

Barnett, W.S., Masse, L.N., 2007a. Comparative benefit–cost analysis of the Abecedarian program and its policy implications. Econ. Educ. Rev., The Economics of Early Childhood Education 26, 113–125. https://doi.org/10.1016/j.econedurev.2005.10.007

Barnett, W.S., Masse, L.N., 2007b. Erratum to: "Comparative benefit–cost analysis of the Abecedarian program and its policy implications": [Econ. Educ. Rev. 26 (2007) 113–125]. Econ. Educ. Rev. 26, 395–396. https://doi.org/10.1016/j.econedurev.2007.02.001

Baron, I.S., 2000. Behavior rating inventory of executive function. Child Neuropsychol. J. Norm. Abnorm. Dev. Child. Adolesc. 6, 235–238. https://doi.org/10.1076/chin.6.3.235.3152

Bastien, P., 2008. R´egression pls et donn´ees censur´ees. Conservatoire National des Arts et M´etiers, Paris.

Amy Johnson – September 2018

Bathelt, J., Gathercole, S.E., Butterfield, S., CALM team, Astle, D.E., 2018a. Children's academic attainment is linked to the global organization of the white matter connectome. Dev. Sci. https://doi.org/10.1111/desc.12662

Bathelt, J., Gathercole, S.E., Johnson, A., Astle, D.E., 2018b. Differences in brain morphology and working memory capacity across childhood. Dev. Sci. 21. https://doi.org/10.1111/desc.12579

Bathelt, J., Holmes, J., Astle, D.E., Centre for Attention Learning and Memory (CALM) Team, 2018c. Data-Driven Subtyping of Executive Function-Related Behavioral Problems in Children. J. Am. Acad. Child Adolesc. Psychiatry 57, 252-262.e4. https://doi.org/10.1016/j.jaac.2018.01.014

Bathelt, J., Scerif, G., Nobre, K., Astle, D., 2018d. Whole-brain white matter organization, intelligence, and educational attainment. bioRxiv 297713. https://doi.org/10.1101/297713

Bavelier, D., Green, C.S., Dye, M.W.G., 2010. Children, wired – for better and for worse. Neuron 67, 692–701. https://doi.org/10.1016/j.neuron.2010.08.035

Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C. a. M., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M., 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. 6, 750–757. https://doi.org/10.1038/nn1075

Bergman Nutley, S., Söderqvist, S., 2017. How Is Working Memory Training Likely to Influence Academic Performance? Current Evidence and Methodological Considerations. Front. Psychol. 8. https://doi.org/10.3389/fpsyg.2017.00069

Best, J.R., Miller, P.H., 2010. A Developmental Perspective on Executive Function. Child Dev. 81, 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bettcher, B.M., Walsh, C.M., Watson, C., Miller, J.W., Green, R., Patel, N., Miller, B.L., Neuhaus, J., Yaffe, K., Kramer, J.H., 2013. Body Mass and White Matter Integrity: The Influence of Vascular and Inflammatory Markers. PLOS ONE 8, e77741. https://doi.org/10.1371/journal.pone.0077741

Biswal, B.B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.-M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kötter, R., Li, S.-J., Lin, C.-P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.A.R.B., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.-J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.-C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.-F., Zhang, H.-Y., Castellanos, F.X., Milham, M.P., 2010. Toward discovery science of human brain function. Proc. Natl. Acad. Sci. U. S. A. 107, 4734–4739. https://doi.org/10.1073/pnas.0911855107

Blanken, L.M.E., White, T., Mous, S.E., Basten, M., Muetzel, R.L., Jaddoe, V.W.V., Wals, M., van der Ende, J., Verhulst, F.C., Tiemeier, H., 2017. Cognitive functioning in children with internalising, externalising and dysregulation problems: a population-based study. Eur. Child Adolesc. Psychiatry 26, 445–456. https://doi.org/10.1007/s00787-016-0903-9

Bokde, A.L.W., Lopez-Bayo, P., Meindl, T., Pechler, S., Born, C., Faltraco, F., Teipel, S.J., Möller, H.-J., Hampel, H., 2006. Functional connectivity of the fusiform gyrus during a face-matching task in subjects with mild cognitive impairment. Brain J. Neurol. 129, 1113–1124. https://doi.org/10.1093/brain/awl051

Bolón-Canedo, V., Sánchez-Maroño, N., Alonso-Betanzos, A., 2015. Recent advances and emerging challenges of feature selection in the context of big data. Knowl.-Based Syst. 86, 33–45. https://doi.org/10.1016/j.knosys.2015.05.014

Boto, E., Meyer, S.S., Shah, V., Alem, O., Knappe, S., Kruger, P., Fromhold, T.M., Lim, M., Glover, P.M., Morris, P.G., Bowtell, R., Barnes, G.R., Brookes, M.J., 2017. A new generation of magnetoencephalography: Room temperature measurements using optically-pumped magnetometers. NeuroImage 149, 404–414. https://doi.org/10.1016/j.neuroimage.2017.01.034

Bradley, R.H., Caldwell, B.M., 1977. Home observation for measurement of the environment: a validation study of screening efficiency. Am. J. Ment. Defic. 81, 417–420.

Bradley, R.H., Caldwell, B.M., Rock, S.L., Barnard, K.E., Gray, C., Hammond, M.A., Mitchell, R., Ramey, C.T., Gottfried, A.W., Johnson, D.L., 1989. Life: A Collaborative Study Involving Six Sites and Three Ethnic Groups in North America. Dev. Psychol. 25, 217–235.

Bradley, R.H., Corwyn, R.F., 2002a. Socioeconomic status and child development. Annu. Rev. Psychol. 53, 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Bradley, R.H., Corwyn, R.F., 2002b. Socioeconomic status and child development. Annu. Rev. Psychol. 53, 371–399. https://doi.org/10.1146/annurev.psych.53.100901.135233

Bradley, R.H., Corwyn, R.F., Burchinal, M., McAdoo, H.P., Coll, C.G., 2001. The Home Environments of Children in the United States Part II: Relations with Behavioral Development through Age Thirteen. Child Dev. 72, 1868–1886. https://doi.org/10.1111/1467-8624.t01-1-00383

Bradley, R.H., Vandell, D.L., 2007. Child care and the well-being of children. Arch. Pediatr. Adolesc. Med. 161, 669–676. https://doi.org/10.1001/archpedi.161.7.669

Bramley, G., Hirsch, D., Littlewood, M., Watkins, D., 2016. Counting the cost of UK poverty. Joseph Rowntree Foundation.

Briggs-Gowan, M.J., Carter, A.S., Skuban, E.M., Horwitz, S.M., 2001. Prevalence of social-emotional and behavioral problems in a community sample of 1- and 2-year-old children. J. Am. Acad. Child Adolesc. Psychiatry 40, 811–819. https://doi.org/10.1097/00004583-200107000-00016

Brito, N.H., Fifer, W.P., Myers, M.M., Elliott, A.J., Noble, K.G., 2016. Associations among family socioeconomic status, EEG power at birth, and cognitive skills during infancy. Dev. Cogn. Neurosci. 19, 144–151. https://doi.org/10.1016/j.dcn.2016.03.004

Brito, N.H., Noble, K.G., 2014. Socioeconomic status and structural brain development. Front. Neurosci. 8, 276. https://doi.org/10.3389/fnins.2014.00276

Brito, N.H., Piccolo, L.R., Noble, K.G., 2017. Associations between cortical thickness and neurocognitive skills during childhood vary by family socioeconomic factors. Brain Cogn. 116, 54–62. https://doi.org/10.1016/j.bandc.2017.03.007

Brookes, M.J., Tewarie, P.K., Hunt, B.A.E., Robson, S.E., Gascoyne, L.E., Liddle, E.B., Liddle, P.F., Morris, P.G., 2016. A multi-layer network approach to MEG connectivity analysis. NeuroImage 132, 425–438. https://doi.org/10.1016/j.neuroimage.2016.02.045

Brookes, M.J., Woolrich, M., Luckhoo, H., Price, D., Hale, J.R., Stephenson, M.C., Barnes, G.R., Smith, S.M., Morris, P.G., 2011. Investigating the electrophysiological basis of resting state networks using

magnetoencephalography. Proc. Natl. Acad. Sci. U. S. A. 108, 16783–16788. https://doi.org/10.1073/pnas.1112685108

Brooks-Gunn, J., Duncan, G.J., 1997a. The effects of poverty on children. Future Child. Cent. Future Child. David Lucile Packard Found. 7, 55–71.

Brooks-Gunn, J., Duncan, G.J., 1997b. The effects of poverty on children. Future Child. Cent. Future Child. David Lucile Packard Found. 7, 55–71.

Butterworth, P., Cherbuin, N., Sachdev, P., Anstey, K.J., 2012a. The association between financial hardship and amygdala and hippocampal volumes: results from the PATH through life project. Soc. Cogn. Affect. Neurosci. 7, 548–556. https://doi.org/10.1093/scan/nsr027

Butterworth, P., Cherbuin, N., Sachdev, P., Anstey, K.J., 2012b. The association between financial hardship and amygdala and hippocampal volumes: results from the PATH through life project. Soc. Cogn. Affect. Neurosci. 7, 548–556. https://doi.org/10.1093/scan/nsr027

Byford, M., Kuh, D., Richards, M., 2012. Parenting practices and intergenerational associations in cognitive ability. Int. J. Epidemiol. 41, 263–272. https://doi.org/10.1093/ije/dyr188

Cain, K., Oakhill, J., Bryant, P., 2004. Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. J. Educ. Psychol. 96, 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Camilli, G., Vargas, S., Ryan, S., Barnett, W.S., 2010. Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development. Teach. Coll. Rec. 112, 579–620.

Cao, M., Wang, J.-H., Dai, Z.-J., Cao, X.-Y., Jiang, L.-L., Fan, F.-M., Song, X.-W., Xia, M.-R., Shu, N., Dong, Q., Milham, M.P., Castellanos, F.X., Zuo, X.-N., He, Y., 2014. Topological organization of the human brain functional connectome across the lifespan. Dev. Cogn. Neurosci. 7, 76–93. https://doi.org/10.1016/j.dcn.2013.11.004

Cao, W., Cao, X., Hou, C., Li, T., Cheng, Y., Jiang, L., Luo, C., Li, C., Yao, D., 2016. Effects of Cognitive Training on Resting-State Functional Connectivity of Default Mode, Salience, and Central Executive Networks. Front. Aging Neurosci. 8. https://doi.org/10.3389/fnagi.2016.00070

Capron, C., Duyme, M., 1989. Assessment of effects of socio-economic status on IQ in a full cross-fostering study. Nature 340, 552–554. https://doi.org/10.1038/340552a0

Case, A., Paxson, C., 2006. Children's health and social mobility. Future Child. 16, 151–173.

Casey, B.J., Giedd, J.N., Thomas, K.M., 2000. Structural and functional brain development and its relation to cognitive development. Biol. Psychol. 54, 241–257.

Casey, B.J., Trainor, R.J., Orendi, J.L., Schubert, A.B., Nystrom, L.E., Giedd, J.N., Castellanos, F.X., Haxby, J.V., Noll, D.C., Cohen, J.D., Forman, S.D., Dahl, R.E., Rapoport, J.L., 1997. A Developmental Functional MRI Study of Prefrontal Activation during Performance of a Go-No-Go Task. J. Cogn. Neurosci. 9, 835–847. https://doi.org/10.1162/jocn.1997.9.6.835

Cattell, R., 1949. Culture Free Intelligence Test. Institute of Personality and Ability Testing, Champaign.

Cattell, R.B., 1963. Theory of fluid and crystallized intelligence: A critical experiment. J. Educ. Psychol. 54, 1–22. https://doi.org/10.1037/h0046743

Cavanagh, J., Krishnadas, R., Batty, G.D., Burns, H., Deans, K.A., Ford, I., McConnachie, A., McGinty, A., McLean, J.S., Millar, K., Sattar, N., Shiels,

P.G., Tannahill, C., Velupillai, Y.N., Packard, C.J., McLean, J., 2013. Socioeconomic Status and the Cerebellar Grey Matter Volume. Data from a Well-Characterised Population Sample. Cerebellum 12, 882–891. https://doi.org/10.1007/s12311-013-0497-4

Chan, M.Y., Na, J., Agres, P.F., Savalia, N.K., Park, D.C., Wig, G.S., 2018. Socioeconomic status moderates age-related differences in the brain's functional network organization and anatomy across the adult lifespan. Proc. Natl. Acad. Sci. 201714021. https://doi.org/10.1073/pnas.1714021115

Chen, Q., Wang, S., 2013. Variable selection for multiply-imputed data with application to dioxin exposure study. Stat. Med. 32, 3646–3659. https://doi.org/10.1002/sim.5783

Cheng, S., Maeda, T., Yoichi, S., Yamagata, Z., Tomiwa, K., 2010. Early Television Exposure and Children's Behavioral and Social Outcomes at Age 30 Months. J. Epidemiol. 20, S482–S489. https://doi.org/10.2188/jea.JE20090179

Chiang, M.-C., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Hickie, I., Toga, A.W., Wright, M.J., Thompson, P.M., 2011. Genetics of white matter development: A DTI study of 705 twins and their siblings aged 12 to 29. NeuroImage 54, 2308–2317. https://doi.org/10.1016/j.neuroimage.2010.10.015

Christakis, D.A., Zimmerman, F.J., DiGiuseppe, D.L., McCarty, C.A., 2004. Early Television Exposure and Subsequent Attentional Problems in Children. Pediatrics 113, 708–713. https://doi.org/10.1542/peds.113.4.708

Chun, H., Keleş, S., 2010. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. J. R. Stat. Soc. Ser. B Stat. Methodol. 72, 3–25. https://doi.org/10.1111/j.1467-9868.2009.00723.x

Colclough, G.L., Brookes, M.J., Smith, S.M., Woolrich, M.W., 2015. A symmetric multivariate leakage correction for MEG connectomes. Neuroimage 117, 439–448. https://doi.org/10.1016/j.neuroimage.2015.03.071

Colclough, G.L., Woolrich, M.W., Tewarie, P.K., Brookes, M.J., Quinn, A.J., Smith, S.M., 2016. How reliable are MEG resting-state connectivity metrics? NeuroImage 138, 284–293. https://doi.org/10.1016/j.neuroimage.2016.05.070

Cole, K., Mitchell, P., 2000. Siblings in the development of executive control and a theory of mind. Br. J. Dev. Psychol. 18, 279–295. https://doi.org/10.1348/026151000165698

Colom, R., Escorial, S., Shih, P.C., Privado, J., 2007. Fluid intelligence, memory span, and temperament difficulties predict academic performance of young adolescents. Personal. Individ. Differ. 42, 1503–1514. https://doi.org/10.1016/j.paid.2006.10.023

Comstock, G.A., Paik, H.-J., 1991. Television and the American Child. Academic Press.

Cooper, K., Stewart, K., 2013. Does money affect children's outcomes? A systematic review. Joseph Rowntree Foundation.

Corak, M., 2006. Do poor children become poor adults? Lessons from a cross country comparison of generational earnings mobility.

Costello, E.J., Compton, S.N., Keeler, G., Angold, A., 2003. Relationships between poverty and psychopathology: a natural experiment. JAMA 290, 2023–2029. https://doi.org/10.1001/jama.290.15.2023

Coupe, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., Barillot, C., 2008. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Trans. Med. Imaging 27, 425–441. https://doi.org/10.1109/TMI.2007.906087

Amy Johnson – September 2018

Craddock, R.C., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2009. Disease state prediction from resting state functional connectivity. Magn. Reson. Med. 62, 1619–1628. https://doi.org/10.1002/mrm.22159

Crook, S.R., Evans, G.W., 2014. The Role of Planning Skills in the Income-Achievement Gap. Child Dev. 85, 405–411. https://doi.org/10.1111/cdev.12129

Czernochowski, D., Fabiani, M., Friedman, D., 2008a. Use it or lose it? SES mitigates age-related decline in a recency/recognition task. Neurobiol. Aging 29, 945–958. https://doi.org/10.1016/j.neurobiolaging.2006.12.017

Czernochowski, D., Fabiani, M., Friedman, D., 2008b. Use it or lose it? SES mitigates age-related decline in a recency/recognition task. Neurobiol. Aging 29, 945–958. https://doi.org/10.1016/j.neurobiolaging.2006.12.017

Dadds, M.R., Maujean, A., Fraser, J.A., 2003. Parenting and conduct problems in children: Australian data and psychometric properties of the alabama parenting questionnaire. Aust. Psychol. 38, 238–241. https://doi.org/10.1080/00050060310001707267

Dai, X., Heckman, J.J., 2013. Older Siblings' Contributions to Young Child's Cognitive Skills. Econ. Model. 35, 235–248. https://doi.org/10.1016/j.econmod.2013.07.003

Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. U. S. A. 103, 13848–13853. https://doi.org/10.1073/pnas.0601417103

D'Angiulli, A., Herdman, A., Stapells, D., Hertzman, C., 2008. Children's event-related potentials of auditory selective attention vary with their socioeconomic status. Neuropsychology 22, 293–300. https://doi.org/10.1037/0894-4105.22.3.293

D'Angiulli, A., Lipina, S.J., Olesinska, A., 2012. Explicit and implicit issues in the developmental cognitive neuroscience of social inequality. Front. Hum. Neurosci. 6. https://doi.org/10.3389/fnhum.2012.00254

Dashiff, C., DiMicco, W., Myers, B., Sheppard, K., 2009. Poverty and adolescent mental health. J. Child Adolesc. Psychiatr. Nurs. Off. Publ. Assoc. Child Adolesc. Psychiatr. Nurses Inc 22, 23–32. https://doi.org/10.1111/j.1744-6171.2008.00166.x

de Moura, L.M., Cogo-Moreira, H., de Ávila, C.R.B., Pan, P.M., Gadelha, A., Moriyama, T., Del Aquilla, M.A., Hoexter, M., Salum, G.A., Picon, F.A., Anés, M., Mercadante, M.T., Lacerda, A., Amaro, E., Miguel, E.C., Rohde, L.A., Bressan, R.A., McGuire, P., Sato, J.R., de Jesus Mari, J., Jackowski, A.P., 2016. Children with Poor Reading Skills at the Word Level Show Reduced Fractional Anisotropy in White Matter Tracts of Both Hemispheres. Brain Connect. 6, 519–523. https://doi.org/10.1089/brain.2016.0430

de Reus, M.A., van den Heuvel, M.P., 2013. Estimating false positives and negatives in brain networks. NeuroImage 70, 402–409. https://doi.org/10.1016/j.neuroimage.2012.12.066

Dearing, E., McCartney, K., Taylor, B.A., 2001. Change in Family Income-to-Needs Matters More for Children with Less. Child Dev. 72, 1779–1793. https://doi.org/10.1111/1467-8624.00378

Deary, I.J., Strand, S., Smith, P., Fernandes, C., 2007. Intelligence and educational achievement. Intelligence 35, 13–21. https://doi.org/10.1016/j.intell.2006.02.001

Decety, J., Yoder, K.J., Lahey, B.B., 2015. Sex differences in abnormal white matter development associated with conduct disorder in children. Psychiatry Res. 233, 269–277. https://doi.org/10.1016/j.pscychresns.2015.07.009

Demerens, C., Stankoff, B., Logak, M., Anglade, P., Allinquant, B., Couraud, F., Zalc, B., Lubetzki, C., 1996. Induction of myelination in the central nervous system by electrical activity. Proc. Natl. Acad. Sci. U. S. A. 93, 9887–9892.

Department for Education, 2017a. Widening participation in higher education: 2017 [WWW Document]. URL https://www.gov.uk/government/statistics/widening-participation-in-higher-education-2017 (accessed 1.20.18).

Department For Education, 2017. A level attainment by pupil characteristics [WWW Document]. URL https://www.gov.uk/government/publications/a-level-attainment-by-pupil-characteristics (accessed 1.20.18).

Department for Education, 2017b. Permanent and fixed-period exclusions in England: 2015 to 2016 [WWW Document]. URL https://www.gov.uk/government/statistics/permanent-and-fixed-period-exclusions-in-england-2015-to-2016 (accessed 1.20.18).

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Deutsch, G.K., Dougherty, R.F., Bammer, R., Siok, W.T., Gabrieli, J.D.E., Wandell, B., 2005. Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. Cortex J. Devoted Study Nerv. Syst. Behav. 41, 354–363.

Devine, R.T., Bignardi, G., Hughes, C., 2016. Executive Function Mediates the Relations between Parental Behaviors and Children's Early Academic Ability. Front. Psychol. 7. https://doi.org/10.3389/fpsyg.2016.01902

Dilworth-Bart, J.E., 2012. Does executive function mediate SES and home quality associations with academic readiness? Early Child. Res. Q. 27, 416–425. https://doi.org/10.1016/j.ecresq.2012.02.002

Donenberg, G., Baker, B.L., 1993. The impact of young children with externalizing behaviors on their families. J. Abnorm. Child Psychol. 21, 179–198.

Dosenbach, N.U.F., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett, J.R., Barch, D.M., Petersen, S.E., Schlaggar, B.L., 2010. Prediction of Individual Brain Maturity Using fMRI. Science 329, 1358–1361. https://doi.org/10.1126/science.1194144

Douet, V., Chang, L., Cloak, C., Ernst, T., 2014. Genetic Influences on Brain Developmental Trajectories on Neuroimaging Studies: From Infancy to Young Adulthood. Brain Imaging Behav. 8, 234–250. https://doi.org/10.1007/s11682-013-9260-1

Downey, D.B., 2001. Number of siblings and intellectual development. The resource dilution explanation. Am. Psychol. 56, 497–504.

Downey, D.B., 1995. When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance. Am. Sociol. Rev. 60, 746–761. https://doi.org/10.2307/2096320

Du, Y., Fu, Z., Calhoun, V.D., 2018. Classification and Prediction of Brain Disorders Using Functional Connectivity: Promising but Challenging. Front. Neurosci. 12. https://doi.org/10.3389/fnins.2018.00525

Dufford, A.J., Kim, P., 2017. Family Income, Cumulative Risk Exposure, and White Matter Structure in Middle Childhood. Front. Hum. Neurosci. 11. https://doi.org/10.3389/fnhum.2017.00547

Duncan, G.J., Brooks-Gunn, J., Klebanov, P.K., 1994. Economic deprivation and early childhood development. Child Dev. 65, 296–318.

Duncan, G.J., Kalil, A., Magnuson, K., Murnane, R., 2014. Two policies to boost school readiness.

Duncan, G.J., Magnuson, K., 2012. Socioeconomic status and cognitive functioning: moving from correlation to causation. Wiley Interdiscip. Rev. Cogn. Sci. 3, 377–386. https://doi.org/10.1002/wcs.1176

Duncan, G.J., Magnuson, K., Votruba-Drzal, E., 2017. Moving Beyond Correlations in Assessing the Consequences of Poverty. Annu. Rev. Psychol. 68, null. https://doi.org/10.1146/annurev-psych-010416-044224

Duncan, O., 1961. A socioeconomic index for all occupations, in: Occupations and Social Status. New York.

Eckert, M.A., Lombardino, L.J., Leonard, C.M., 2001. Planar asymmetry tips the phonological playground and environment raises the bar. Child Dev. 72, 988–1002.

Education Endowment Foundation, 2018. Teaching and Learning Toolkit [WWW Document]. Educ. Endow. Found. URL https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit (accessed 1.20.18).

Education standards research team, 2012. Research evidence on reading for pleasure. Department For Education.

Eide, E.R., Showalter, M.H., Goldhaber, D.D., 2010. The relation between children's health and academic achievement. Child. Youth Serv. Rev. 32, 231–238. https://doi.org/10.1016/j.childyouth.2009.08.019

El Nokali, N.E., Bachman, H.J., Votruba-Drzal, E., 2010. Parent Involvement and Children's Academic and Social Development in Elementary School. Child Dev. 81, 988–1005. https://doi.org/10.1111/j.1467-8624.2010.01447.x

Elardo, R., Bradley, R.H., 1981. The home observation for measurement of the environment (HOME) scale: A review of research. Dev. Rev. 1, 113–145. https://doi.org/10.1016/0273-2297(81)90012-5

Eluvathingal, T.J., Chugani, H.T., Behen, M.E., Juhász, C., Muzik, O., Maqbool, M., Chugani, D.C., Makki, M., 2006. Abnormal brain connectivity in children after early severe socioemotional deprivation: a diffusion tensor imaging study. Pediatrics 117, 2093–2100. https://doi.org/10.1542/peds.2005-1727

End Child Poverty, 2015. Why End Child Poverty [WWW Document]. End Child Poverty. URL http://endchildpoverty.org.uk/ (accessed 6.11.15).

Engle, P.L., Black, M.M., Behrman, J.R., Cabral de Mello, M., Gertler, P.J., Kapiriri, L., Martorell, R., Young, M.E., 2007. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. The Lancet 369, 229–242. https://doi.org/10.1016/S0140-6736(07)60112-3

Evans, G.W., 2004. The environment of childhood poverty. Am. Psychol. 59, 77–92. https://doi.org/10.1037/0003-066X.59.2.77

Evans, G.W., Kim, P., 2007. Childhood poverty and health: cumulative risk exposure and stress dysregulation. Psychol. Sci. 18, 953–957. https://doi.org/10.1111/j.1467-9280.2007.02008.x

Evans, G.W., Schamberg, M.A., 2009. Childhood poverty, chronic stress, and adult working memory. Proc. Natl. Acad. Sci. 106, 6545–6549. https://doi.org/10.1073/pnas.0811910106

Evans, T.M., Kochalka, J., Ngoon, T.J., Wu, S.S., Qin, S., Battista, C., Menon, V., 2015. Brain Structural Integrity and Intrinsic Functional Connectivity Forecast 6

Year Longitudinal Growth in Children's Numerical Abilities. J. Neurosci. 35, 11743–11750. https://doi.org/10.1523/JNEUROSCI.0216-15.2015

Fair, D.A., Bathula, D., Nikolas, M.A., Nigg, J.T., 2012. Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. Proc. Natl. Acad. Sci. 109, 6769–6774. https://doi.org/10.1073/pnas.1115365109

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U.F., Church, J.A., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2009. Functional Brain Networks Develop from a "Local to Distributed" Organization. PLoS Comput. Biol. 5. https://doi.org/10.1371/journal.pcbi.1000381

Farah, M.J., 2017. The Neuroscience of Socioeconomic Status: Correlates, Causes, and Consequences. Neuron 96, 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

Farah, M.J., Shera, D.M., Savage, J.H., Betancourt, L., Giannetta, J.M., Brodsky, N.L., Malmud, E.K., Hurt, H., 2006. Childhood poverty: Specific associations with neurocognitive development. Brain Res. 1110, 166–174. https://doi.org/10.1016/j.brainres.2006.06.072

Featherman, D.L., Lancaster Jones, F., Hauser, R.M., 1975. Assumptions of social mobility research in the U.S.: The case of occupational status. Soc. Sci. Res. 4, 329–360. https://doi.org/10.1016/0049-089X(75)90002-2

Feinstein, L., 2000. The relative economic importance of academic, psychological and behavioural attributes developed on childhood [WWW Document]. URL http://cep.lse.ac.uk (accessed 7.2.18).

Fernald, A., Marchman, V.A., Weisleder, A., 2013. SES differences in language processing skill and vocabulary are evident at 18 months. Dev. Sci. 16, 234–248. https://doi.org/10.1111/desc.12019

Fields, R.D., 2008. White matter in learning, cognition and psychiatric disorders. Trends Neurosci. 31, 361–370. https://doi.org/10.1016/j.tins.2008.04.001

Finch, W.H., 2016. Missing Data and Multiple Imputation in the Context of Multivariate Analysis of Variance. J. Exp. Educ. 84, 356–372. https://doi.org/10.1080/00220973.2015.1011594

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18, 1664–1671. https://doi.org/10.1038/nn.4135

Flöel, A., de Vries, M.H., Scholz, J., Breitenstein, C., Johansen-Berg, H., 2009. White matter integrity in the vicinity of Broca's area predicts grammar learning success. NeuroImage 47, 1974–1981. https://doi.org/10.1016/j.neuroimage.2009.05.046

Flouri, E., Midouhas, E., Joshi, H., 2015. Family and neighbourhood risk and children's problem behaviour: The moderating role of intelligence. Intelligence 53, 33–42. https://doi.org/10.1016/j.intell.2015.08.003

Fornito, A., Zalesky, A., Pantelis, C., Bullmore, E.T., 2012. Schizophrenia, neuroimaging and connectomics. NeuroImage 62, 2296–2314. https://doi.org/10.1016/j.neuroimage.2011.12.090

Frederickson, N., Reason, R., 1997. Phonological Assessment Battery (PhAB): Manual and Test Materials. NFER-Nelson.

Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9, 432–441. https://doi.org/10.1093/biostatistics/kxm045

Garali, I., Adanyeguh, I.M., Ichou, F., Perlbarg, V., Seyer, A., Colsch, B., Moszer, I., Guillemot, V., Durr, A., Mochel, F., Tenenhaus, A., 2017. A strategy for

multimodal data integration: application to biomarkers identification in spinocerebellar ataxia. Brief. Bioinform. https://doi.org/10.1093/bib/bbx060

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., van der Walt, S., Descoteaux, M., Nimmo-Smith, I., Dipy Contributors, 2014. Dipy, a library for the analysis of diffusion MRI data. Front. Neuroinformatics 8, 8. https://doi.org/10.3389/fninf.2014.00008

Gianaros, P.J., Horenstein, J.A., Hariri, A.R., Sheu, L.K., Manuck, S.B., Matthews, K.A., Cohen, S., 2008. Potential neural embedding of parental social standing. Soc. Cogn. Affect. Neurosci. 3, 91–96. https://doi.org/10.1093/scan/nsn003

Gianaros, P.J., Marsland, A.L., Sheu, L.K., Erickson, K.I., Verstynen, T.D., 2013. Inflammatory pathways link socioeconomic inequalities to white matter architecture. Cereb. Cortex N. Y. N 1991 23, 2058–2071. https://doi.org/10.1093/cercor/bhs191

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. Nat. Neurosci. 2, 861–863. https://doi.org/10.1038/13158

Ginkel, J.R. van, Kroonenberg, P.M., 2014. Using Generalized Procrustes Analysis for Multiple Imputation in Principal Component Analysis. J. Classif. 31, 242–269. https://doi.org/10.1007/s00357-014-9154-y

Giorgio, A., Watkins, K.E., Chadwick, M., James, S., Winmill, L., Douaud, G., De Stefano, N., Matthews, P.M., Smith, S.M., Johansen-Berg, H., James, A.C., 2010. Longitudinal changes in grey and white matter during adolescence. NeuroImage 49, 94–103. https://doi.org/10.1016/j.neuroimage.2009.08.003

Giorgio, A., Watkins, K.E., Douaud, G., James, A.C., James, S., De Stefano, N., Matthews, P.M., Smith, S.M., Johansen-Berg, H., 2008. Changes in white matter microstructure during adolescence. NeuroImage 39, 52–61. https://doi.org/10.1016/j.neuroimage.2007.07.043

Goodall, C., 1991. Procrustes Methods in the Statistical Analysis of Shape. J. R. Stat. Soc. Ser. B Methodol. 53, 285–339.

Goodman, R., 1997. The Strengths and Difficulties Questionnaire: a research note. J. Child Psychol. Psychiatry 38, 581–586.

Gorman, K., 2010. The Consequences of Multicollinearity among Socioeconomic Predictors of Negative Concord in Philadelphia. Univ. Pa. Work. Pap. Linguist. 16.

Gregg, P., Propper, C., Washbrook, E., 2007. Understanding the Relationship between Parental Income and Multiple Child Outcomes: A Decomposition Analysis (SSRN Scholarly Paper No. ID 1159360). Social Science Research Network, Rochester, NY.

Greicius, M.D., Menon, V., 2004. Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. J. Cogn. Neurosci. 16, 1484–1492. https://doi.org/10.1162/0898929042568532

Griggs, J., Walker, R., 2008. The costs of child poverty for individuals and society: a literature review. Joseph Rowntree Foundation.

Guerra-Carrillo, B., Mackey, A.P., Bunge, S.A., 2014. Resting-State fMRI: A Window into Human Brain Plasticity. The Neuroscientist 20, 522–533. https://doi.org/10.1177/1073858414524442

Gullick, M.M., Demir-Lira, Ö.E., Booth, J.R., 2016. Reading skill–fractional anisotropy relationships in visuospatial tracts diverge depending on socioeconomic status. Dev. Sci. 19, 673–685. https://doi.org/10.1111/desc.12428

Gunst, R.F., Mason, R.L., 1979. Some Considerations in the Evaluation of Alternate Prediction Equations. Technometrics 21, 55–63. https://doi.org/10.2307/1268580

Günther, O.P., Shin, H., Ng, R.T., McMaster, W.R., McManus, B.M., Keown, P.A., Tebbutt, S.J., Lê Cao, K.-A., 2014. Novel multivariate methods for integration of genomics and proteomics data: applications in a kidney transplant rejection study. Omics J. Integr. Biol. 18, 682–695. https://doi.org/10.1089/omi.2014.0062

Hackman, D.A., Farah, M.J., 2009. Socioeconomic status and the developing brain. Trends Cogn. Sci. 13, 65–73. https://doi.org/10.1016/j.tics.2008.11.003

Hackman, D.A., Farah, M.J., Meaney, M.J., 2010. Socioeconomic status and the brain: mechanistic insights from human and animal research. Nat. Rev. Neurosci. 11, 651–659. https://doi.org/10.1038/nrn2897

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O., 2008. Mapping the Structural Core of Human Cerebral Cortex. PLOS Biol. 6, e159. https://doi.org/10.1371/journal.pbio.0060159

Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V.J., Meuli, R., Thiran, J.-P., 2007. Mapping human whole-brain structural networks with diffusion MRI. PloS One 2, e597. https://doi.org/10.1371/journal.pone.0000597

Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V.J., Meuli, R., Thiran, J.-P., Grant, P.E., 2010. White matter maturation reshapes structural connectivity in the late developing human brain. Proc. Natl. Acad. Sci. U. S. A. 107, 19067–19072. https://doi.org/10.1073/pnas.1009073107

Hair, J.F., Ringle, C.M., Sarstedt, M., 2011. PLS-SEM: Indeed a Silver Bullet. J. Mark. Theory Pract. 19, 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, N.L., Hanson, J.L., Wolfe, B.L., Pollak, S.D., 2015. Association of Child Poverty, Brain Development, and Academic Achievement. JAMA Pediatr. 169, 822–829. https://doi.org/10.1001/jamapediatrics.2015.1475

Hamadani, J.D., Tofail, F., Huda, S.N., Alam, D.S., Ridout, D.A., Attanasio, O., Grantham-McGregor, S.M., 2014. Cognitive Deficit and Poverty in the First 5 Years of Childhood in Bangladesh. Pediatrics. https://doi.org/10.1542/peds.2014-0694

Hampson, M., Driesen, N.R., Skudlarski, P., Gore, J.C., Constable, R.T., 2006a. Brain Connectivity Related to Working Memory Performance. J. Neurosci. 26, 13338–13343. https://doi.org/10.1523/JNEUROSCI.3408-06.2006

Hampson, M., Tokoglu, F., Sun, Z., Schafer, R.J., Skudlarski, P., Gore, J.C., Constable, R.T., 2006b. Connectivity-behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. NeuroImage 31, 513–519. https://doi.org/10.1016/j.neuroimage.2005.12.040

Hanscombe, K.B., Trzaskowski, M., Haworth, C.M.A., Davis, O.S.P., Dale, P.S., Plomin, R., 2012. Socioeconomic status (SES) and children's intelligence (IQ): in a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. PloS One 7, e30320. https://doi.org/10.1371/journal.pone.0030320

Hanson, J.L., Chandra, A., Wolfe, B.L., Pollak, S.D., 2011a. Association between Income and the Hippocampus. PLoS ONE 6. https://doi.org/10.1371/journal.pone.0018712

Hanson, J.L., Chandra, A., Wolfe, B.L., Pollak, S.D., 2011b. Association between Income and the Hippocampus. PLoS ONE 6. https://doi.org/10.1371/journal.pone.0018712

Hanson, J.L., Chung, M.K., Avants, B.B., Rudolph, K.D., Shirtcliff, E.A., Gee, J.C., Davidson, R.J., Pollak, S.D., 2012. Structural variations in prefrontal cortex

mediate the relationship between early childhood stress and spatial working memory. J. Neurosci. 32, 7917–7925. https://doi.org/10.1523/JNEUROSCI.0307-12.2012

Hanson, J.L., Hair, N., Shen, D.G., Shi, F., Gilmore, J.H., Wolfe, B.L., Pollak, S.D., 2013. Family Poverty Affects the Rate of Human Infant Brain Growth. PLoS ONE 8. https://doi.org/10.1371/journal.pone.0080954

Harpur, L.J., Polek, E., van Harmelen, A.-L., 2015. The role of timing of maltreatment and child intelligence in pathways to low symptoms of depression and anxiety in adolescence. Child Abuse Negl. 47, 24–37. https://doi.org/10.1016/j.chiabu.2015.05.019

Harris, C., Straker, L., Pollock, C., 2017. A socioeconomic related "digital divide" exists in how, not if, young people use computers. PLOS ONE 12, e0175011. https://doi.org/10.1371/journal.pone.0175011

Hastings, E.C., Karas, T.L., Winsler, A., Way, E., Madigan, A., Tyler, S., 2009. Young Children's Video/Computer Game Use: Relations with School Performance and Behavior. Issues Ment. Health Nurs. 30, 638–649. https://doi.org/10.1080/01612840903050414

Hauser, R.M., Warren, J.R., 1997. Socioeconomic Indexes for Occupations: A Review, Update, and Critique. Sociol. Methodol. 27, 177–298. https://doi.org/10.1111/1467-9531.271028

Hawkins, E., Astle, D.E., Rastle, K., 2015. Semantic advantage for learning new phonological form representations. J. Cogn. Neurosci. 27, 775–786. https://doi.org/10.1162/jocn_a_00730

Hecker, T., Hermenau, K., Salmen, C., Teicher, M., Elbert, T., 2016. Harsh discipline relates to internalizing problems and cognitive functioning: findings from a cross-sectional study with school children in Tanzania. BMC Psychiatry 16. https://doi.org/10.1186/s12888-016-0828-3

Heckman, J.J., Stixrud, J., Urzua, S., 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior (Working Paper No. 12006). National Bureau of Economic Research. https://doi.org/10.3386/w12006

Heiland, F., 2009. Does the birth order affect the cognitive development of a child? Appl. Econ. 41, 1799–1818. https://doi.org/10.1080/00036840601083220

Helton, J., Vaughn, M., Kavish, N., Boutwell, B.B., 2018. The Association of Externalizing and Internalizing Problems with Indicators of Intelligence in a Sample of At-Risk Children. https://doi.org/10.1101/210500

Henseler, J., Hubona, G., Ray, P.A., 2016. Using PLS path modeling in new technology research: updated guidelines. Ind. Manag. Data Syst. 116, 2–20. https://doi.org/10.1108/IMDS-09-2015-0382

Hermundstad, A.M., Bassett, D.S., Brown, K.S., Aminoff, E.M., Clewett, D., Freeman, S., Frithsen, A., Johnson, A., Tipper, C.M., Miller, M.B., Grafton, S.T., Carlson, J.M., 2013. Structural foundations of resting-state and task-based functional connectivity in the human brain. Proc. Natl. Acad. Sci. 110, 6169–6174. https://doi.org/10.1073/pnas.1219562110

Herrod, H.G., 2007. Do first years really last a lifetime? Clin. Pediatr. (Phila.) 46, 199–205. https://doi.org/10.1177/0009922806297303

Hillebrand, A., Barnes, G.R., 2002. A Quantitative Assessment of the Sensitivity of Whole-Head MEG to Activity in the Adult Human Cortex. NeuroImage 16, 638–650. https://doi.org/10.1006/nimg.2002.1102

Hinshaw, S.P., 1992. Externalizing behavior problems and academic underachievement in childhood and adolescence: causal relationships and underlying mechanisms. Psychol. Bull. 111, 127–155.

Hipp, J.F., Hawellek, D.J., Corbetta, M., Siegel, M., Engel, A.K., 2012. Large-scale cortical correlation structure of spontaneous oscillatory activity. Nat. Neurosci. 15, 884–890. https://doi.org/10.1038/nn.3101

Hoeft, F., McCandliss, B.D., Black, J.M., Gantman, A., Zakerani, N., Hulme, C., Lyytinen, H., Whitfield-Gabrieli, S., Glover, G.H., Reiss, A.L., Gabrieli, J.D.E., 2011. Neural systems predicting long-term outcome in dyslexia. Proc. Natl. Acad. Sci. U. S. A. 108, 361–366. https://doi.org/10.1073/pnas.1008950108

Hofer, S.M., Clouston, S., 2014. Commentary: On the Importance of Early Life Cognitive Abilities in Shaping Later Life Outcomes. Res. Hum. Dev. 11, 241–246. https://doi.org/10.1080/15427609.2014.936173

Hollingshead, A.B., 1975. Four factor index of social status.

Holmes, T.H., Rahe, R.H., 1967. The social readjustment rating scale. J. Psychosom. Res. 11, 213–218. https://doi.org/10.1016/0022-3999(67)90010-4

Holmgren, S., Molander, B., Nilsson, L.-G., 2006. Intelligence and executive functioning in adult age: Effects of sibship size and birth order. Eur. J. Cogn. Psychol. 18, 138–158. https://doi.org/10.1080/09541440500216150

Holz, N.E., Boecker, R., Hohm, E., Zohsel, K., Buchmann, A.F., Blomeyer, D., Jennen-Steinmetz, C., Baumeister, S., Hohmann, S., Wolf, I., Plichta, M.M., Esser, G., Schmidt, M., Meyer-Lindenberg, A., Banaschewski, T., Brandeis, D., Laucht, M., 2015. The Long-Term Impact of Early Life Poverty on Orbitofrontal Cortex Volume in Adulthood: Results from a Prospective Study Over 25 Years. Neuropsychopharmacology 40, 996. https://doi.org/10.1038/npp.2014.277

Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc. Natl. Acad. Sci. U. S. A. 106, 2035–2040. https://doi.org/10.1073/pnas.0811168106

Howard-Jones, P., 2014. Neuroscience and Education: A Review of Educational Interventions and Approaches Informed by Neuroscience.

Howes, C., 1988. Relations Between Early Child Care and Schooling. Dev. Psychol. 24.

Huang, J., Zhang, T., 2009. The Benefit of Group Sparsity. ArXiv09012962 Math Stat.

Huber, B., Yeates, M., Meyer, D., Fleckhammer, L., Kaufman, J., 2018. The effects of screen media content on young children's executive functioning. J. Exp. Child Psychol. 170, 72–85. https://doi.org/10.1016/j.jecp.2018.01.006

Hughes, C., Devine, R.T., 2017. For Better or for Worse? Positive and Negative Parental Influences on Young Children's Executive Function. Child Dev. https://doi.org/10.1111/cdev.12915

Hughes, C., Ensor, R., 2011. Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. J. Exp. Child Psychol. 108, 663–676. https://doi.org/10.1016/j.jecp.2010.06.005

Hull, J.V., Dokovna, L.B., Jacokes, Z.J., Torgerson, C.M., Irimia, A., Van Horn, J.D., 2016. Resting-State Functional Connectivity in Autism Spectrum Disorders: A Review. Front. Psychiatry 7, 205. https://doi.org/10.3389/fpsyt.2016.00205

Huttenlocher, P.R., Dabholkar, A.S., 1997. Regional differences in synaptogenesis in human cerebral cortex. J. Comp. Neurol. 387, 167–178.

Hwang, K., Hallquist, M.N., Luna, B., 2013. The Development of Hub Architecture in the Human Functional Brain Network. Cereb. Cortex N. Y. NY 23, 2380–2393. https://doi.org/10.1093/cercor/bhs227

Ikuta, T., Shafritz, K.M., Bregman, J., Peters, B., Gruner, P., Malhotra, A.K., Szeszko, P.R., 2014. Abnormal Cingulum Bundle Development in Autism: A

Probabilistic Tractography Study. Psychiatry Res. 221, 63–68. https://doi.org/10.1016/j.pscychresns.2013.08.002

Iltus, S., 2006. Significance of home environments as proxy indicators for early childhood care and education, Education for All Global Monitoring Report 2007; Strong foundations: early childhood care and education.

Ingoldsby, E.M., Shaw, D.S., 2002. Neighborhood contextual factors and early-starting antisocial pathways. Clin. Child Fam. Psychol. Rev. 5, 21–55.

Institute for Social and Economic Research, 2015. Understanding Society – The UK Household Longitudinal Study [WWW Document]. Underst. Soc. URL https://www.understandingsociety.ac.uk/ (accessed 6.8.17).

Jednorog, K., Altarelli, I., Monzalvo, K., Fluss, J., Dubois, J., Billard, C., Dehaene-Lambertz, G., Ramus, F., 2012. The Influence of Socioeconomic Status on Children's Brain Structure. PLoS ONE 7. https://doi.org/10.1371/journal.pone.0042486

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. NeuroImage 62, 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Johansen-Berg, H., 2010. Behavioural relevance of variation in white matter microstructure. Curr. Opin. Neurol. 23, 351–358. https://doi.org/10.1097/WCO.0b013e32833b7631

Johnson, M.H., 2011. Interactive Specialization: A domain-general framework for human functional brain development? Dev. Cogn. Neurosci. 1, 7–21. https://doi.org/10.1016/j.dcn.2010.07.003

Johnson, S.B., Riis, J.L., Noble, K.G., 2016. State of the Art Review: Poverty and the Developing Brain. Pediatrics 137. https://doi.org/10.1542/peds.2015-3075

Jolles, D., Ashkenazi, S., Kochalka, J., Evans, T., Richardson, J., Rosenberg-Lee, M., Zhao, H., Supekar, K., Chen, T., Menon, V., 2016a. Parietal hyper-connectivity, aberrant brain organization, and circuit-based biomarkers in children with mathematical disabilities. Dev. Sci. 19, 613–631. https://doi.org/10.1111/desc.12399

Jolles, D., Supekar, K., Richardson, J., Tenison, C., Ashkenazi, S., Rosenberg-Lee, M., Fuchs, L., Menon, V., 2016b. Reconfiguration of parietal circuits with cognitive tutoring in elementary school children. Cortex 83, 231–245. https://doi.org/10.1016/j.cortex.2016.08.004

Jung, M., Tu, Y., Lang, C.A., Ortiz, A., Park, J., Jorgenson, K., Kong, X.-J., Kong, J., 2017. Decreased structural connectivity and resting-state brain activity in the lateral occipital cortex is associated with social communication deficits in boys with autism spectrum disorder. NeuroImage. https://doi.org/10.1016/j.neuroimage.2017.09.031

Kaiser, M., 2017. Mechanisms of Connectome Development. Trends Cogn. Sci. 21, 703–717. https://doi.org/10.1016/j.tics.2017.05.010

Kamijo, K., Pontifex, M.B., O'Leary, K.C., Scudder, M.R., Wu, C.-T., Castelli, D.M., Hillman, C.H., 2011. The effects of an afterschool physical activity program on working memory in preadolescent children. Dev. Sci. 14, 1046–1058. https://doi.org/10.1111/j.1467-7687.2011.01054.x

Kaminski, M., Blinowska, K.J., 2014. Directed Transfer Function is not influenced by volume conduction-inexpedient pre-processing should be avoided. Front. Comput. Neurosci. 8, 61. https://doi.org/10.3389/fncom.2014.00061

Karwath, C., Relikowski, I., Schmitt, M., 2014. Sibling structure and educational achievement: How do the number of siblings, birth order, and birth spacing affect children's vocabulary competences? ZfF – Z. Für Fam. J. Fam. Res. 26.

Keating, D.P., 2013. Cognitive and Brain Development, in: Handbook of Adolescent Psychology. Wiley-Blackwell, pp. 45–84. https://doi.org/10.1002/9780471726746.ch3

Keller, T.A., Just, M.A., 2009. Altering cortical connectivity: remediation-induced changes in the white matter of poor readers. Neuron 64, 624–631. https://doi.org/10.1016/j.neuron.2009.10.018

Kim, E.M., Sheridan, S.M., Kwon, K., Koziol, N., 2013. Parent beliefs and children's social-behavioral functioning: The mediating role of parent–teacher relationships. J. Sch. Psychol. 51, 175–185. https://doi.org/10.1016/j.jsp.2013.01.003

Kim, P., Evans, G.W., Angstadt, M., Ho, S.S., Sripada, C.S., Swain, J.E., Liberzon, I., Phan, K.L., 2013. Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. Proc. Natl. Acad. Sci. U. S. A. 110, 18442–18447. https://doi.org/10.1073/pnas.1308240110

Kishiyama, M.M., Boyce, W.T., Jimenez, A.M., Perry, L.M., Knight, R.T., 2009. Socioeconomic disparities affect prefrontal function in children. J. Cogn. Neurosci. 21, 1106–1115. https://doi.org/10.1162/jocn.2009.21101

Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J.D., Moseley, M.E., Poldrack, R.A., 2000. Microstructure of temporo-parietal white matter as a basis for reading ability: evidence from diffusion tensor magnetic resonance imaging. Neuron 25, 493–500.

Kobayashi, T., Kawachi, I., Iwase, T., Suzuki, E., Takao, S., 2013. Individual-level social capital and self-rated health in Japan: An application of the Resource Generator. Soc. Sci. Med. 85, 32–37. https://doi.org/10.1016/j.socscimed.2013.02.027

Koyama, M.S., Di Martino, A., Zuo, X.-N., Kelly, C., Mennes, M., Jutagir, D.R., Castellanos, F.X., Milham, M.P., 2011. Resting-state functional connectivity indexes reading competence in children and adults. J. Neurosci. Off. J. Soc. Neurosci. 31, 8617–8624. https://doi.org/10.1523/JNEUROSCI.4865-10.2011

Koyama, M.S., Martino, A.D., Kelly, C., Jutagir, D.R., Sunshine, J., Schwartz, S.J., Castellanos, F.X., Milham, M.P., 2013. Cortical Signatures of Dyslexia and Remediation: An Intrinsic Functional Connectivity Approach. PLOS ONE 8, e55454. https://doi.org/10.1371/journal.pone.0055454

Kremer, P., Elshaug, C., Leslie, E., Toumbourou, J.W., Patton, G.C., Williams, J., 2014. Physical activity, leisure-time screen use and depression among children and young adolescents. J. Sci. Med. Sport 17, 183–187. https://doi.org/10.1016/j.jsams.2013.03.012

Kring, S.R., Greenberg, J.S., Seltzer, M.M., 2010. The Impact of Health Problems on Behavior Problems in Adolescents and Adults With Autism Spectrum Disorders: Implications for Maternal Burden. Soc. Work Ment. Health 8, 54–71. https://doi.org/10.1080/15332980902932441

Krishnadas, R., Kim, J., McLean, J., Batty, G.D., McLean, J.S., Millar, K., Packard, C.J., Cavanagh, J., 2013. The envirome and the connectome: exploring the structural noise in the human brain associated with socioeconomic deprivation. Front. Hum. Neurosci. 7. https://doi.org/10.3389/fnhum.2013.00722

Krishnan, A., Williams, L.J., McIntosh, A.R., Abdi, H., 2011. Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. NeuroImage 56, 455–475. https://doi.org/10.1016/j.neuroimage.2010.07.034

Lange, N., Froimowitz, M.P., Bigler, E.D., Lainhart, J.E., 2010. Associations Between IQ, Total and Regional Brain Volumes and Demography in a Large Normative

Sample of Healthy Children and Adolescents. Dev. Neuropsychol. 35, 296–317. https://doi.org/10.1080/87565641003696833

Lauder, J.M., 1983. Hormonal and humoral influences on brain development. Psychoneuroendocrinology 8, 121–155.

Lawson, G.M., Duda, J.T., Avants, B.B., Wu, J., Farah, M.J., 2013. Associations between children's socioeconomic status and prefrontal cortical thickness. Dev. Sci. 16, 641–652. https://doi.org/10.1111/desc.12096

Lawson, G.M., Farah, M.J., 2017. Executive function as a mediator between SES and academic achievement throughout childhood   , Executive function as a mediator between SES and academic achievement throughout childhood. Int. J. Behav. Dev. 41, 94–104. https://doi.org/10.1177/0165025415603489

Le Floch, É., Guillemot, V., Frouin, V., Pinel, P., Lalanne, C., Trinchera, L., Tenenhaus, A., Moreno, A., Zilbovicius, M., Bourgeron, T., Dehaene, S., Thirion, B., Poline, J.-B., Duchesnay, É., 2012. Significant correlation between a set of genetic polymorphisms and a functional brain network revealed by feature selection and sparse Partial Least Squares. NeuroImage 63, 11–24. https://doi.org/10.1016/j.neuroimage.2012.06.061

Lebel, C., Treit, S., Beaulieu, C., 2017. A review of diffusion MRI of typical white matter development from early childhood to young adulthood. NMR Biomed. https://doi.org/10.1002/nbm.3778

Leventhal, T., Brooks-Gunn, J., 2000. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. Psychol. Bull. 126, 309–337.

Li, C., Yuan, H., Urbano, D., Cha, Y., Ding, L., 2017. ICA on sensor or source data: A comparison study in deriving resting state networks from EEG, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Presented at the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3604–3607. https://doi.org/10.1109/EMBC.2017.8037637

Li, W., Farkas, G., Duncan, G.J., Burchinal, M.R., Vandell, D.L., 2013. Timing of High-Quality Child Care and Cognitive, Language, and Preacademic Development. Dev. Psychol. 49, 1440–1451. https://doi.org/10.1037/a0030613

Licht, C., 2010. New methods for generating significance levels from multiply-imputed data. Otto-Friedrich-University Bamberg.

Lillard, A.S., Peterson, J., 2011. The Immediate Impact of Different Types of Television on Young Children's Executive Function. Pediatrics 128, 644–649. https://doi.org/10.1542/peds.2010-1919

Lipina, S., Segretin, S., Hermida, J., Prats, L., Fracchia, C., Camelo, J.L., Colombo, J., 2013. Linking childhood poverty and cognition: environmental mediators of non-verbal executive control in an Argentine sample. Dev. Sci. 16, 697–707. https://doi.org/10.1111/desc.12080

Lipina, S.J., 2016. Critical considerations about the use of poverty measures in the study of cognitive development. Int. J. Psychol. n/a-n/a. https://doi.org/10.1002/ijop.12282

Lipina, S.J., Evers, K., 2017. Neuroscience of Childhood Poverty: Evidence of Impacts and Mechanisms as Vehicles of Dialog With Ethics. Front. Psychol. 8. https://doi.org/10.3389/fpsyg.2017.00061

Lipina, S.J., Martelli, M.I., Vuelta, B., Colombo, J.A., 2005. Performance on the A-not-B task of Argentinean infants from unsatisfied and satisfied basic needs homes. Interam. J. Psychol. 39, 49.

Lipina, S.J., Posner, M.I., 2012. The impact of poverty on the development of brain networks. Front. Hum. Neurosci. 6, 238. https://doi.org/10.3389/fnhum.2012.00238

Llewellyn, D.J., Lang, I.A., Langa, K.M., Huppert, F.A., 2008. Cognitive function and psychological well-being: findings from a population-based cohort. Age Ageing 37, 685–689. https://doi.org/10.1093/ageing/afn194

Loe, I.M., Lee, E.S., Feldman, H.M., 2013. Attention and Internalizing Behaviors in Relation to White Matter in Children Born Preterm. J. Dev. Behav. Pediatr. JDBP 34, 156–164. https://doi.org/10.1097/DBP.0b013e3182842122

Loeber, R., Stouthamer-Loeber, M., 1986. Family Factors as Correlates and Predictors of Juvenile Conduct Problems and Delinquency. Crime Justice 7, 29–149.

Luby, J., Belden, A., Botteron, K., Marrus, N., Harms, M.P., Babb, C., Nishino, T., Barch, D., 2013. The Effects of Poverty on Childhood Brain Development: The Mediating Effect of Caregiving and Stressful Life Events. JAMA Pediatr. 167, 1135–1142. https://doi.org/10.1001/jamapediatrics.2013.3139

Luckhoo, H., Hale, J.R., Stokes, M.G., Nobre, A.C., Morris, P.G., Brookes, M.J., Woolrich, M.W., 2012. Inferring task-related networks using independent component analysis in magnetoencephalography. NeuroImage 62, 530–541. https://doi.org/10.1016/j.neuroimage.2012.04.046

Luo, C., Tu, S., Peng, Y., Gao, S., Li, J., Dong, L., Li, G., Lai, Y., Li, H., Yao, D., 2014. Long-term effects of musical training and functional plasticity in salience system. Neural Plast. 2014, 180138. https://doi.org/10.1155/2014/180138

Lupie, S.J., King, S., Meaney, M.J., McEwen, B.S., 2001. Can poverty get under your skin? basal cortisol levels and cognitive function in children from low and high socioeconomic status. Dev. Psychopathol. 13, 653–676.

Mabbott, D.J., Noseworthy, M., Bouffet, E., Laughlin, S., Rockel, C., 2006. White matter growth as a mechanism of cognitive development in children. NeuroImage 33, 936–946. https://doi.org/10.1016/j.neuroimage.2006.07.024

Mackey, A.P., Miller Singley, A.T., Bunge, S.A., 2013. Intensive reasoning training alters patterns of brain connectivity at rest. J. Neurosci. Off. J. Soc. Neurosci. 33, 4796–4803. https://doi.org/10.1523/JNEUROSCI.4141-12.2013

MacKinnon, D.P., Warsi, G., Dwyer, J.H., 1995. A Simulation Study of Mediated Effect Measures. Multivar. Behav. Res. 30, 41. https://doi.org/10.1207/s15327906mbr3001_3

Marcus Jenkins, J.V., Woolley, D.P., Hooper, S.R., De Bellis, M.D., 2013. Direct and Indirect Effects of Brain Volume, Socioeconomic Status and Family Stress on Child IQ. J. Child Adolesc. Behav. 1.

Marrelec, G., Krainik, A., Duffau, H., Pélégrini-Issac, M., Lehéricy, S., Doyon, J., Benali, H., 2006. Partial correlation for functional brain interactivity investigation in functional MRI. NeuroImage 32, 228–237. https://doi.org/10.1016/j.neuroimage.2005.12.057

Masten, A.S., Hubbard, J.J., Gest, S.D., Tellegen, A., Garmezy, N., Ramirez, M., 1999. Competence in the context of adversity: pathways to resilience and maladaptation from childhood to late adolescence. Dev. Psychopathol. 11, 143–169.

Matejko, A.A., Ansari, D., 2015. Drawing connections between white matter and numerical and mathematical cognition: A literature review. Neurosci. Biobehav. Rev. 48, 35–52. https://doi.org/10.1016/j.neubiorev.2014.11.006

Matthews, M., Fair, D.A., 2015. Research Review: Functional brain connectivity and child psychopathology – overview and methodological considerations for

investigators new to the field. J. Child Psychol. Psychiatry 56, 400–414. https://doi.org/10.1111/jcpp.12335

McClelland, M.M., Acock, A.C., Piccinin, A., Rhea, S.A., Stallings, M.C., 2013. Relations between Preschool Attention Span-Persistence and Age 25 Educational Outcomes. Early Child. Res. Q. 28, 314–324. https://doi.org/10.1016/j.ecresq.2012.07.008

McDowell, K.D., Lonigan, C.J., Goldstein, H., 2007. Relations Among Socioeconomic Status, Age, and Predictors of Phonological Awareness. J. Speech Lang. Hear. Res. 50, 1079. https://doi.org/10.1044/1092-4388(2007/075)

McGrew, K.S., Woodcock, R.W., Schrank, K.A., 2007. Woodcock-Johnson III normative update technical manual. Riverside Pub., Itasca, IL.

McIntosh, A.R., Lobaugh, N.J., 2004. Partial least squares analysis of neuroimaging data: applications and advances. NeuroImage, Mathematics in Brain Imaging 23, S250–S263. https://doi.org/10.1016/j.neuroimage.2004.07.020

McLeod, J.D., Shanahan, M.J., 1993. Poverty, Parenting, and Children's Mental Health. Am. Sociol. Rev. 58, 351–366. https://doi.org/10.2307/2095905

McLoyd, V.C., 1998. Socioeconomic disadvantage and child development. Am. Psychol. 53, 185–204.

McNab, F., Zeidman, P., Rutledge, R.B., Smittenaar, P., Brown, H.R., Adams, R.A., Dolan, R.J., 2015. Age-related changes in working memory and the ability to ignore distraction. Proc. Natl. Acad. Sci. 112, 6515–6518. https://doi.org/10.1073/pnas.1504162112

Meng, C., Zeleznik, O.A., Thallinger, G.G., Kuster, B., Gholami, A.M., Culhane, A.C., 2016. Dimension reduction techniques for the integrative analysis of multi-omics data. Brief. Bioinform. 17, 628–641. https://doi.org/10.1093/bib/bbv108

Menting, B., van Lier, P.A.C., Koot, H.M., 2011. Language skills, peer rejection, and the development of externalizing behavior from kindergarten to fourth grade. J. Child Psychol. Psychiatry 52, 72–79. https://doi.org/10.1111/j.1469-7610.2010.02279.x

Mevel, K., Fransson, P., Bölte, S., 2015. Multimodal brain imaging in autism spectrum disorder and the promise of twin research. Autism Int. J. Res. Pract. 19, 527–541. https://doi.org/10.1177/1362361314535510

Mezzacappa, E., 2004. Alerting, orienting, and executive attention: developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. Child Dev. 75, 1373–1386. https://doi.org/10.1111/j.1467-8624.2004.00746.x

Milligan, K., Stabile, M., 2011. Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions. Am. Econ. J. Econ. Policy 3, 175–205. https://doi.org/10.1257/pol.3.3.175

Mistry, K.B., Minkovitz, C.S., Strobino, D.M., Borzekowski, D.L.G., 2007. Children's television exposure and behavioral and social outcomes at 5.5 years: does timing of exposure matter? Pediatrics 120, 762–769. https://doi.org/10.1542/peds.2006-3573

Monteiro, J.M., Rao, A., Shawe-Taylor, J., Mourão-Miranda, J., 2016. A multiple hold-out framework for Sparse Partial Least Squares. J. Neurosci. Methods 271, 182–194. https://doi.org/10.1016/j.jneumeth.2016.06.011

Moons, K.G.M., Donders, R.A.R.T., Stijnen, T., Harrell, F.E., 2006. Using the outcome for imputation of missing predictor values was preferred. J. Clin. Epidemiol. 59, 1092–1101. https://doi.org/10.1016/j.jclinepi.2006.01.009

Mosher, J.C., Leahy, R.M., Lewis, P.S., 1999. EEG and MEG: Forward Solutions for Inverse Methods. IEEE Trans. Biomed. Eng. 46, 15.

Muetzel, R.L., Blanken, L.M.E., van der Ende, J., El Marroun, H., Shaw, P., Sudre, G., van der Lugt, A., Jaddoe, V.W.V., Verhulst, F.C., Tiemeier, H., White, T., 2018. Tracking Brain Development and Dimensional Psychiatric Symptoms in Children: A Longitudinal Population-Based Neuroimaging Study. Am. J. Psychiatry 175, 54–62. https://doi.org/10.1176/appi.ajp.2017.16070813

Muftuler, L.T., Davis, E.P., Buss, C., Solodkin, A., Su, M.Y., Head, K.M., Hasso, A.N., Sandman, C.A., 2012. Development of white matter pathways in typically developing preadolescent children. Brain Res. 1466, 33–43. https://doi.org/10.1016/j.brainres.2012.05.035

Muscatell, K.A., Morelli, S.A., Falk, E.B., Way, B.M., Pfeifer, J.H., Galinsky, A.D., Lieberman, M.D., Dapretto, M., Eisenberger, N.I., 2012. Social status modulates neural activity in the mentalizing network. NeuroImage 60, 1771–1777. https://doi.org/10.1016/j.neuroimage.2012.01.080

Nadler, B., Coifman, R.R., 2005. The prediction error in CLS and PLS: the importance of feature selection prior to multivariate calibration. J. Chemom. 19, 107–118. https://doi.org/10.1002/cem.915

Nagy, Z., Westerberg, H., Klingberg, T., 2004. Maturation of white matter is associated with the development of cognitive functions during childhood. J. Cogn. Neurosci. 16, 1227–1233. https://doi.org/10.1162/0898929041920441

Najman, J.M., Clavarino, A., McGee, T.R., Bor, W., Williams, G.M., Hayatbakhsh, M.R., 2010a. Timing and Chronicity of Family Poverty and Development of Unhealthy Behaviors in Children: A Longitudinal Study. J. Adolesc. Health 46, 538–544. https://doi.org/10.1016/j.jadohealth.2009.12.001

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

Nathanson, A.I., Aladé, F., Sharp, M.L., Rasmussen, E.E., Christy, K., 2014. The relation between television exposure and executive function among preschoolers. Dev. Psychol. 50, 1497–1506. https://doi.org/10.1037/a0035714

Neece, C.L., Green, S.A., Baker, B.L., 2012. Parenting Stress and Child Behavior Problems: A Transactional Relationship Across Time. Am. J. Intellect. Dev. Disabil. 117, 48–66. https://doi.org/10.1352/1944-7558-117.1.48

Nesbitt, K.T., Baker-Ward, L., Willoughby, M.T., 2013. Executive function mediates socio-economic and racial differences in early academic achievement. Early Child. Res. Q. 28, 774–783. https://doi.org/10.1016/j.ecresq.2013.07.005

Neville, H.J., Stevens, C., Pakulak, E., Bell, T.A., Fanning, J., Klein, S., Isbell, E., 2013. Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. Proc. Natl. Acad. Sci. 110, 12138–12143. https://doi.org/10.1073/pnas.1304437110

NICHD Early Child Care Research Network, 2003. Do children's attention processes mediate the link between family predictors and school readiness? Dev. Psychol. 39, 581–593. https://doi.org/10.1037/0012-1649.39.3.581

Nitzl, C., Roldan, J.L., Cepeda, G., 2016. Mediation analysis in partial least squares path modeling: Helping researchers discuss more sophisticated models. Ind. Manag. Data Syst. 116, 1849–1864. https://doi.org/10.1108/IMDS-07-2015-0302

Noble, K.G., Farah, M.J., McCandliss, B.D., 2006a. Socioeconomic background modulates cognition–achievement relationships in reading. Cogn. Dev. 21, 349–368. https://doi.org/10.1016/j.cogdev.2006.01.007

Noble, K.G., Grieve, S.M., Korgaonkar, M.S., Engelhardt, L.E., Griffith, E.Y., Williams, L.M., Brickman, A.M., 2012a. Hippocampal volume varies with educational attainment across the life-span. Front. Hum. Neurosci. 6, 307. https://doi.org/10.3389/fnhum.2012.00307

Noble, K.G., Houston, S.M., Brito, N.H., Bartsch, H., Kan, E., Kuperman, J.M., Akshoomoff, N., Amaral, D.G., Bloss, C.S., Libiger, O., Schork, N.J., Murray, S.S., Casey, B.J., Chang, L., Ernst, T.M., Frazier, J.A., Gruen, J.R., Kennedy, D.N., Van Zijl, P., Mostofsky, S., Kaufmann, W.E., Kenet, T., Dale, A.M., Jernigan, T.L., Sowell, E.R., 2015. Family income, parental education and brain structure in children and adolescents. Nat. Neurosci. advance online publication. https://doi.org/10.1038/nn.3983

Noble, K.G., Houston, S.M., Kan, E., Sowell, E.R., 2012b. Neural correlates of socioeconomic status in the developing human brain. Dev. Sci. 15, 516–527. https://doi.org/10.1111/j.1467-7687.2012.01147.x

Noble, K.G., Houston, S.M., Kan, E., Sowell, E.R., 2012c. Neural correlates of socioeconomic status in the developing human brain. Dev. Sci. 15, 516–527. https://doi.org/10.1111/j.1467-7687.2012.01147.x

Noble, K.G., Korgaonkar, M.S., Grieve, S.M., Brickman, A.M., 2013. Higher Education is an Age-Independent Predictor of White Matter Integrity and Cognitive Control in Late Adolescence. Dev. Sci. 16, 653–664. https://doi.org/10.1111/desc.12077

Noble, K.G., McCandliss, B.D., Farah, M.J., 2007. Socioeconomic gradients predict individual differences in neurocognitive abilities. Dev. Sci. 10, 464–480. https://doi.org/10.1111/j.1467-7687.2007.00600.x

Noble, K.G., Norman, M.F., Farah, M.J., 2005. Neurocognitive correlates of socioeconomic status in kindergarten children. Dev. Sci. 8, 74–87. https://doi.org/10.1111/j.1467-7687.2005.00394.x

Noble, K.G., Wolmetz, M.E., Ochs, L.G., Farah, M.J., McCandliss, B.D., 2006b. Brain–behavior relationships in reading acquisition are modulated by socioeconomic factors. Dev. Sci. 9, 642–654. https://doi.org/10.1111/j.1467-7687.2006.00542.x

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., Hallett, M., 2004. Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 115, 2292–2307. https://doi.org/10.1016/j.clinph.2004.04.029

Nores, M., Barnett, W.S., 2010. Benefits of early childhood interventions across the world: (Under) Investing in the very young. Econ. Educ. Rev., Special Issue in Honor of Henry M. Levin 29, 271–282. https://doi.org/10.1016/j.econedurev.2009.09.001

Nunes, A.S., Peatfield, N., Vakorin, V., Doesburg, S.M., 2018. Idiosyncratic organization of cortical networks in autism spectrum disorder. NeuroImage. https://doi.org/10.1016/j.neuroimage.2018.01.022

Oakhill, J.V., Cain, K., 2012. The Precursors of Reading Ability in Young Readers: Evidence From a Four-Year Longitudinal Study. Sci. Stud. Read. 16, 91–121. https://doi.org/10.1080/10888438.2010.529219

OECD, 2002. Reading for Change: Performance and Engagement across Countries, PISA. OECD Publishing. https://doi.org/10.1787/9789264099289-en

Office for National Statistics, 2016a. National pupil database [WWW Document]. URL https://www.gov.uk/government/collections/national-pupil-database (accessed 1.10.18).

Office for National Statistics, 2016b. The National Statistics Socio-economic classification (NS-SEC) [WWW Document]. URL

https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifica
tions/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010
(accessed 6.8.17).

Office for National Statistics, E.L. and S.A., 2016. SOC2010 volume 2: the structure
and coding index [WWW Document]. Off. Natl. Stat. URL
https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccup
ationalclassificationsoc/soc2010/soc2010volume2thestructureandcodingindex
(accessed 6.8.17).

Office for National Statistics, M. of H., Communities &.Local Government, 2015.
English indices of deprivation 2015 [WWW Document]. GOV.UK. URL
https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015
(accessed 9.24.18).

OHBA Analysis Group [WWW Document], 2017. . GitHub. URL
https://github.com/OHBA-analysis (accessed 9.24.18).

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open source
software for advanced analysis of MEG, EEG, and invasive electrophysiological
data. Comput. Intell. Neurosci. 2011, 156869.
https://doi.org/10.1155/2011/156869

Otero, G.A., 1997. Poverty, cultural disadvantage and brain development: a study of
pre-school children in Mexico. Electroencephalogr. Clin. Neurophysiol. 102,
512–516.

Pages, J., 2004. Analyse factorielle de donnees mixtes. Revue Statistique Appliquee LII,
93–111.

Pampel, F.C., Krueger, P.M., Denney, J.T., 2010. Socioeconomic Disparities in Health
Behaviors. Annu. Rev. Sociol. 36, 349–370.
https://doi.org/10.1146/annurev.soc.012809.102529

Pang, E.W., Iii, S., C, O., 2016. From Structure to Circuits: The Contribution of MEG
Connectivity Studies to Functional Neurosurgery. Front. Neuroanat. 10.
https://doi.org/10.3389/fnana.2016.00067

Parkes, A., Sweeting, H., Wight, D., Henderson, M., 2013. Do television and electronic
games predict children's psychosocial adjustment? Longitudinal research using
the UK Millennium Cohort Study. Arch. Dis. Child. 98, 341–348.
https://doi.org/Parkes, A. <http://eprints.gla.ac.uk/view/author/2512.html>,
Sweeting, H. <http://eprints.gla.ac.uk/view/author/6296.html>, Wight, D.
<http://eprints.gla.ac.uk/view/author/4250.html> and Henderson, M.
<http://eprints.gla.ac.uk/view/author/9886.html> (2013) Do television and
electronic games predict children's psychosocial adjustment? Longitudinal
research using the UK Millennium Cohort Study.   Archives of Disease in
Childhood
<http://eprints.gla.ac.uk/view/journal_volume/Archives_of_Disease_in_Childho
od.html>, 98(5), pp. 341-348. (doi:10.1136/archdischild-2011-301508
<http://dx.doi.org/10.1136/archdischild-2011-301508>)

Passolunghi, M.C., Vercelloni, B., Schadee, H., 2007. The precursors of mathematics
learning: Working memory, phonological ability and numerical competence.
Cogn. Dev. 22, 165–184. https://doi.org/10.1016/j.cogdev.2006.09.001

Patalay, P., Fitzsimons, E., 2016. Correlates of Mental Illness and Wellbeing in
Children: Are They the Same? Results From the UK Millennium Cohort Study.
J. Am. Acad. Child Adolesc. Psychiatry 55, 771–783.
https://doi.org/10.1016/j.jaac.2016.05.019

Pennington, B.F., Ozonoff, S., 1996. Executive functions and developmental
psychopathology. J. Child Psychol. Psychiatry 37, 51–87.

Penny, W.D., Friston, K.J., Ashburner, J.T., Kiebel, S.J., Nichols, T.E., 2011. Statistical Parametric Mapping: The Analysis of Functional Brain Images. Elsevier.

Price, G.R., Yeo, D.J., Wilkey, E.D., Cutting, L.E., 2018. Prospective relations between resting-state connectivity of parietal subdivisions and arithmetic competence. Dev. Cogn. Neurosci. 30, 280–290. https://doi.org/10.1016/j.dcn.2017.02.006

Prime, H., Pauker, S., Plamondon, A., Perlman, M., Jenkins, J., 2014. Sibship Size, Sibling Cognitive Sensitivity, and Children's Receptive Vocabulary. Pediatrics 133, e394–e401. https://doi.org/10.1542/peds.2012-2874

Pu, W., Luo, Q., Jiang, Y., Gao, Y., Ming, Q., Yao, S., 2017. Alterations of Brain Functional Architecture Associated with Psychopathic Traits in Male Adolescents with Conduct Disorder. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-11775-z

Pugh, K.R., Mencl, W.E., Jenner, A.R., Katz, L., Frost, S.J., Lee, J.R., Shaywitz, S.E., Shaywitz, B.A., 2001. Neurobiological studies of reading and reading disability. J. Commun. Disord. 34, 479–492.

Puma, M., Bell, S., Cook, R., Heid, C., 2010. Head Start Impact Study: Final Report. Head Start Impact Study, Washington, D.C.

Qiu, D., Tan, L.-H., Zhou, K., Khong, P.-L., 2008. Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. NeuroImage 41, 223–232. https://doi.org/10.1016/j.neuroimage.2008.02.023

Raghubar, K.P., Barnes, M.A., Hecht, S.A., 2010. Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. Learn. Individ. Differ., Perspectives on Math Difficulty and Disability in Children 20, 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Raizada, R.D.S., Kishiyama, M.M., 2010. Effects of Socioeconomic Status on Brain Development, and How Cognitive Neuroscience May Contribute to Levelling the Playing Field. Front. Hum. Neurosci. 4. https://doi.org/10.3389/neuro.09.003.2010

Raizada, R.D.S., Richards, T.L., Meltzoff, A., Kuhl, P.K., 2008. Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. NeuroImage 40, 1392–1401. https://doi.org/10.1016/j.neuroimage.2008.01.021

Reiss, F., 2013. Socioeconomic inequalities and mental health problems in children and adolescents: A systematic review. Soc. Sci. Med. 90, 24–31. https://doi.org/10.1016/j.socscimed.2013.04.026

Richards, M., Stephen, A., Mishra, G., 2010. Health returns to cognitive capital in the British 1946 birth cohort. Longitud. Life Course Stud. 1, 281–296. https://doi.org/10.14301/llcs.v1i3.94

Richmond, S., Johnson, K.A., Seal, M.L., Allen, N.B., Whittle, S., 2016. Development of brain networks and relevance of environmental and genetic factors: A systematic review. Neurosci. Biobehav. Rev. 71, 215–239. https://doi.org/10.1016/j.neubiorev.2016.08.024

Roberts, R.E., Anderson, E.J., Husain, M., 2013. White Matter Microstructure and Cognitive Function. The Neuroscientist 19, 8–15. https://doi.org/10.1177/1073858411421218

Rohart, F., Gautier, B., Singh, A., Lê Cao, K.-A., 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13, e1005752. https://doi.org/10.1371/journal.pcbi.1005752

Rohde, T.E., Thompson, L.A., 2007. Predicting academic achievement with cognitive ability. Intelligence 35, 83–92. https://doi.org/10.1016/j.intell.2006.05.004

Rombouts, S., Scheltens, P., 2005. Functional connectivity in elderly controls and AD patients using resting state fMRI: a pilot study. Curr. Alzheimer Res. 2, 115–116.

Rose, D., Pevalin, D.J., 2003. A researcher's guide to the national statistics socio-economic classification. SAGE.

Rubin, D.B., 2004. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52, 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

Sala, G., Gobet, F., 2017. Does Far Transfer Exist? Negative Evidence From Chess, Music, and Working Memory Training. Curr. Dir. Psychol. Sci. 26, 515–520. https://doi.org/10.1177/0963721417712760

Sampson, R.J., Morenoff, J.D., Gannon-Rowley, T., 2002. Assessing "Neighborhood Effects": Social Processes and New Directions in Research. Annu. Rev. Sociol. 28, 443–478. https://doi.org/10.1146/annurev.soc.28.110601.141114

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., Boyce, W.T., 2011. Family Socioeconomic Status and Child Executive Functions: The Roles of Language, Home Environment, and Single Parenthood. J. Int. Neuropsychol. Soc. 17, 120–132. https://doi.org/10.1017/S1355617710001335

Schmidt, A., Diwadkar, V.A., Smieskova, R., Harrisberger, F., Lang, U.E., McGuire, P., Fusar-Poli, P., Borgwardt, S., 2014. Approaching a network connectivity-driven classification of the psychosis continuum: a selective review and suggestions for future research. Front. Hum. Neurosci. 8, 1047. https://doi.org/10.3389/fnhum.2014.01047

Schmidt, M.E., Rich, M., Rifas-Shiman, S.L., Oken, E., Taveras, E.M., 2009. Television Viewing in Infancy and Child Cognition at 3 Years of Age in a US Cohort. Pediatrics 123, e370–e375. https://doi.org/10.1542/peds.2008-3221

Schmidt, M.E., Vandewater, E.A., 2008. Media and Attention, Cognition, and School Achievement. Future Child. 18, 63–85. https://doi.org/10.1353/foc.0.0004

Schneiders, J., Drukker, M., Ende, J. van der, Verhulst, F.C., Os, J. van, Nicolson, N.A., 2003. Neighbourhood socioeconomic disadvantage and behavioural problems from late childhood into early adolescence. J. Epidemiol. Community Health 57, 699–703. https://doi.org/10.1136/jech.57.9.699

Scholz, J., Tomassini, V., Johansen-Berg, H., 2009. Individual Differences in White Matter Microstructure in the Healthy Brain, in: Behrens, T., Johansen-Berg, H. (Eds.), Diffusion MRI: From Quantitative Measurement to In Vivo Neuroanatomy. Elsevier, Oxford, pp. 237–250.

Scottish Government, 2005. SEED Sponsored Research: Children starting school in Scotland [WWW Document]. URL http://www.gov.scot/Publications/2005/02/20634/51605 (accessed 6.15.15).

Seigneuric, A., Ehrlich, M.-F., 2005. Contribution of working memory capacity to children's reading comprehension: A longitudinal investigation. Read. Writ. 18, 617–656. https://doi.org/10.1007/s11145-005-2038-0

Shaw, B., Bernardes, E., Trethewey, A., Menzies, L., 2016. Special educational needs and their links to poverty. Joseph Rowntree Foundation.

Shehzad, Z., Kelly, A.M.C., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X.,

Milham, M.P., 2009. The resting brain: unconstrained yet reliable. Cereb. Cortex N. Y. N 1991 19, 2209–2229. https://doi.org/10.1093/cercor/bhn256

Sheridan, M.A., Sarsour, K., Jutte, D., D'Esposito, M., Boyce, W.T., 2012. The Impact of Social Disparity on Prefrontal Function in Childhood. PLoS ONE 7, e35744. https://doi.org/10.1371/journal.pone.0035744

Sibley, B.A., Etnier, J.L., 2003. The Relationship between Physical Activity and Cognition in Children: A Meta-Analysis. Pediatr. Exerc. Sci. 15, 243–256. https://doi.org/10.1123/pes.15.3.243

Singer, M.I., Flannery, D.J., Guo, S., Miller, D., Leibbrandt, S., 2004. Exposure to violence, parental monitoring, and television viewing as contributors to children's psychological trauma. J. Community Psychol. 32, 489–504. https://doi.org/10.1002/jcop.20015

Singer, M.I., Slovak, K., Frierson, T., York, P., 1998. Viewing Preferences, Symptoms of Psychological Trauma, and Violent Behaviors Among Children Who Watch Television. J. Am. Acad. Child Adolesc. Psychiatry 37, 1041–1048. https://doi.org/10.1097/00004583-199810000-00014

Skeen, S., Macedo, A., Tomlinson, M., Hensels, I.S., Sherr, L., 2016. Exposure to violence and psychological well-being over time in children affected by HIV/AIDS in South Africa and Malawi. AIDS Care 28, 16–25. https://doi.org/10.1080/09540121.2016.1146219

Smilde, A.K., van der Werf, M.J., Bijlsma, S., van der Werff-van der Vat, B.J.C., Jellema, R.H., 2005. Fusion of mass spectrometry-based metabolomics data. Anal. Chem. 77, 6729–6736. https://doi.org/10.1021/ac051080y

Smith, N., Middleton, S., 2007. Poverty dynamics research in the UK. Joseph Rowntree Foundation.

Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W., 2011. Network modelling methods for FMRI. NeuroImage 54, 875–891. https://doi.org/10.1016/j.neuroimage.2010.08.063

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E.J., Glasser, M.F., Ugurbil, K., Barch, D.M., Van Essen, D.C., Miller, K.L., 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat. Neurosci. 18, 1565–1567. https://doi.org/10.1038/nn.4125

Snook, L., Paulson, L.-A., Roy, D., Phillips, L., Beaulieu, C., 2005. Diffusion tensor imaging of neurodevelopment in children and young adults. NeuroImage 26, 1164–1173. https://doi.org/10.1016/j.neuroimage.2005.03.016

Sorhaindo, A., Feinstein, L., Centre for Research on the Wider Benefits of Learning (London, E., 2006. What is the relationship between child nutrition and school outcomes? (No. 18), Wider Benefits of Learning Research Report. Centre for Research on the Wider Benefits of Learning, London.

Sporns, O., 2013. Structure and function of complex brain networks. Dialogues Clin. Neurosci. 15, 247–262.

Sripada, R.K., Swain, J.E., Evans, G.W., Welsh, R.C., Liberzon, I., 2014. Childhood poverty and stress reactivity are associated with aberrant functional connectivity in default mode network. Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol. 39, 2244–2251. https://doi.org/10.1038/npp.2014.75

St Clair-Thompson, H.L., Gathercole, S.E., 2006. Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. Q. J. Exp. Psychol. 59, 745–759. https://doi.org/10.1080/17470210500162854

Staff, R.T., Murray, A.D., Ahearn, T.S., Mustafa, N., Fox, H.C., Whalley, L.J., 2012. Childhood socioeconomic status and adult brain size: childhood socioeconomic

status influences adult hippocampal size. Ann. Neurol. 71, 653–660. https://doi.org/10.1002/ana.22631

Stam, C.J., Nolte, G., Daffertshofer, A., 2007. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193. https://doi.org/10.1002/hbm.20346

Steinberg, L., 2005. Cognitive and affective development in adolescence. Trends Cogn. Sci. 9, 69–74. https://doi.org/10.1016/j.tics.2004.12.005

Stekhoven, D.J., Bühlmann, P., 2012. MissForest - nonparametric missing value imputation for mixed-type data. Bioinformatics 28, 112–118. https://doi.org/10.1093/bioinformatics/btr597

Stephen Gorard, Huat See, B., Peter, D., 2012. The impact of attitudes and aspirations on educational attainment and participation. Joseph Rowntree Foundation.

Stevens, A.A., Tappon, S.C., Garg, A., Fair, D.A., 2012. Functional Brain Network Modularity Captures Inter- and Intra-Individual Variation in Working Memory Capacity. PLOS ONE 7, e30468. https://doi.org/10.1371/journal.pone.0030468

Stevens, C., Lauinger, B., Neville, H., 2009. Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: An event-related brain potential study. Dev. Sci. 12, 634–646. https://doi.org/10.1111/j.1467-7687.2009.00807.x

Stevens, G., Featherman, D.L., 1981. A revised socioeconomic index of occupational status. Soc. Sci. Res. 10, 364–395. https://doi.org/10.1016/0049-089X(81)90011-9

Strohschein, L., 2005. Household Income Histories and Child Mental Health Trajectories. J. Health Soc. Behav. 46, 359–375. https://doi.org/10.1177/002214650504600404

Strohschein, L., Gauthier, A.H., 2017. Poverty Dynamics, Parenting, and Child Mental Health in Canada. Soc. Ment. Health 2156869317731603. https://doi.org/10.1177/2156869317731603

Sullivan, A., Brown, M., 2013. Social inequalities in cognitive scores at age 16: The role of reading [WWW Document]. CLS Work. Pap. URL http://www.cls.ioe.ac.uk/page.aspx?sitesectionid=939 (accessed 6.11.18).

Supekar, K., Swigart, A.G., Tenison, C., Jolles, D.D., Rosenberg-Lee, M., Fuchs, L., Menon, V., 2013. Neural predictors of individual differences in response to math tutoring in primary-grade school children. Proc. Natl. Acad. Sci. 110, 8230–8235. https://doi.org/10.1073/pnas.1222154110

Supekar, K., Uddin, L.Q., Prater, K., Amin, H., Greicius, M.D., Menon, V., 2010. Development of functional and structural connectivity within the default mode network in young children. NeuroImage 52, 290–301. https://doi.org/10.1016/j.neuroimage.2010.04.009

Swanson, H.L., Jerman, O., 2007. The influence of working memory on reading growth in subgroups of children with reading disabilities. J. Exp. Child Psychol. 96, 249–283. https://doi.org/10.1016/j.jecp.2006.12.004

Swing, E.L., Gentile, D.A., Anderson, C.A., Walsh, D.A., 2010. Television and Video Game Exposure and the Development of Attention Problems. Pediatrics 126, 214–221. https://doi.org/10.1542/peds.2009-1508

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., Yamanouchi, T., Suzuki, S., Kawashima, R., 2010. Training of Working Memory Impacts Structural Connectivity. J. Neurosci. 30, 3297–3303. https://doi.org/10.1523/JNEUROSCI.4611-09.2010

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., Kawashima, R., 2011. Verbal working memory performance correlates with regional white matter structures in the frontoparietal regions. Neuropsychologia 49, 3466–3473. https://doi.org/10.1016/j.neuropsychologia.2011.08.022

Tang, Y.-Y., Lu, Q., Geng, X., Stein, E.A., Yang, Y., Posner, M.I., 2010. Short-term meditation induces white matter changes in the anterior cingulate. Proc. Natl. Acad. Sci. U. S. A. 107, 15649–15652. https://doi.org/10.1073/pnas.1011043107

Taren, A.A., Gianaros, P.J., Greco, C.M., Lindsay, E.K., Fairgrieve, A., Brown, K.W., Rosen, R.K., Ferris, J.L., Julson, E., Marsland, A.L., Bursley, J.K., Ramsburg, J., Creswell, J.D., 2015. Mindfulness meditation training alters stress-related amygdala resting state functional connectivity: a randomized controlled trial. Soc. Cogn. Affect. Neurosci. 10, 1758–1768. https://doi.org/10.1093/scan/nsv066

Tau, G.Z., Peterson, B.S., 2010. Normal Development of Brain Circuits. Neuropsychopharmacology 35, 147–168. https://doi.org/10.1038/npp.2009.115

Taylor, A.B., MacKinnon, D.P., 2012. Four applications of permutation methods to testing a single-mediator model. Behav. Res. Methods 44, 806–844. https://doi.org/10.3758/s13428-011-0181-x

Taylor, S.E., Eisenberger, N.I., Saxbe, D., Lehman, B.J., Lieberman, M.D., 2006. Neural responses to emotional stimuli are associated with childhood family stress. Biol. Psychiatry 60, 296–301. https://doi.org/10.1016/j.biopsych.2005.09.027

Teale, W.H., 1981. Parents Reading to Their Children: What We Know and Need to Know. Lang. Arts 58, 902–912.

Teffer, K., Semendeferi, K., 2012. Chapter 9 - Human prefrontal cortex: Evolution, development, and pathology, in: Hofman, M.A., Falk, D. (Eds.), Progress in Brain Research, Evolution of the Primate Brain. Elsevier, pp. 191–218. https://doi.org/10.1016/B978-0-444-53860-4.00009-X

Tenenhaus, A., Tenenhaus, M., 2014. Regularized generalized canonical correlation analysis for multiblock or multigroup data analysis. Eur. J. Oper. Res. 238, 391–403. https://doi.org/10.1016/j.ejor.2014.01.008

Tenenhaus, A., Tenenhaus, M., 2011. Regularized Generalized Canonical Correlation Analysis. Psychometrika 76, 257. https://doi.org/10.1007/s11336-011-9206-8

Tenenhaus, M., Vinzi, V.E., Chatelin, Y.-M., Lauro, C., 2005. PLS path modeling. Comput. Stat. Data Anal., Partial Least Squares 48, 159–205. https://doi.org/10.1016/j.csda.2004.03.005

The Early Head Start Research Consortium, 2002. Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start Volume I: Final Technical Report. Early Head Start Research and Evaluation Project, Washington, DC.

The MathWorks Inc., 2012. MATLAB. The MathWorks Inc., Massachusetts.

The Reading Agency, 2015. Literature Review: The impact of reading for pleasure and empowerment. BOP Consulting.

Tian, L., Jiang, T., Wang, Y., Zang, Y., He, Y., Liang, M., Sui, M., Cao, Q., Hu, S., Peng, M., Zhuo, Y., 2006. Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder. Neurosci. Lett. 400, 39–43. https://doi.org/10.1016/j.neulet.2006.02.022

Tibshirani, R., 1994. Regression Shrinkage and Selection Via the Lasso. J. R. Stat. Soc. Ser. B 58, 267–288.

Tomalski, P., Moore, D.G., Ribeiro, H., Axelsson, E.L., Murphy, E., Karmiloff-Smith, A., Johnson, M.H., Kushnerenko, E., 2013. Socioeconomic status and functional brain development – associations in early infancy. Dev. Sci. 16, 676–687. https://doi.org/10.1111/desc.12079

Tomarken, A.J., Dichter, G.S., Garber, J., Simien, C., 2004. Resting frontal brain activity: linkages to maternal depression and socio-economic status among adolescents. Biol. Psychol. 67, 77–102. https://doi.org/10.1016/j.biopsycho.2004.03.011

Tomporowski, P.D., McCullick, B., Pendleton, D.M., Pesce, C., 2015. Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. J. Sport Health Sci. 4, 47–55. https://doi.org/10.1016/j.jshs.2014.09.003

Tononi, G., Edelman, G.M., 2000. Schizophrenia and the mechanisms of conscious integration. Brain Res. Brain Res. Rev. 31, 391–400.

Toplak, M.E., West, R.F., Stanovich, K.E., 2013. Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? J. Child Psychol. Psychiatry 54, 131–143. https://doi.org/10.1111/jcpp.12001

Totsika, V., Sylva, K., 2004. The Home Observation for Measurement of the Environment Revisited. Child Adolesc. Ment. Health 9, 25–35. https://doi.org/10.1046/j.1475-357X.2003.00073.x

Tournier, J.-D., Yeh, C.-H., Calamante, F., Cho, K.-H., Connelly, A., Lin, C.-P., 2008. Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. NeuroImage 42, 617–625. https://doi.org/10.1016/j.neuroimage.2008.05.002

Tsang, J.M., Dougherty, R.F., Deutsch, G.K., Wandell, B.A., Ben-Shachar, M., 2009. Frontoparietal white matter diffusion properties predict mental arithmetic skills in children. Proc. Natl. Acad. Sci. U. S. A. 106, 22546–22551. https://doi.org/10.1073/pnas.0906094106

Tuch, D.S., Salat, D.H., Wisco, J.J., Zaleta, A.K., Hevelone, N.D., Rosas, H.D., 2005. Choice reaction time performance correlates with diffusion anisotropy in white matter pathways supporting visuospatial attention. Proc. Natl. Acad. Sci. U. S. A. 102, 12212–12217. https://doi.org/10.1073/pnas.0407259102

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., Gottesman, I.I., 2003. Socioeconomic status modifies heritability of IQ in young children. Psychol. Sci. 14, 623–628.

Uddin, L.Q., Supekar, K., Menon, V., 2013. Reconceptualizing functional brain connectivity in autism from a developmental perspective. Front. Hum. Neurosci. 7, 458. https://doi.org/10.3389/fnhum.2013.00458

Uddin, L.Q., Supekar, K.S., Ryali, S., Menon, V., 2011. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. J. Neurosci. Off. J. Soc. Neurosci. 31, 18578–18589. https://doi.org/10.1523/JNEUROSCI.4465-11.2011

UNICEF Office of Research, 2017. Building the Future: Children and the Sustainable Development Goals in Rich Countries (No. 14). UNICEF.

Ursache, A., Noble, K.G., 2016. Neurocognitive development in socioeconomic context: Multiple mechanisms and implications for measuring socioeconomic status. Psychophysiology 53, 71–82. https://doi.org/10.1111/psyp.12547

Ursache, A., Noble, K.G., Pediatric Imaging, Neurocognition and Genetics Study, 2016. Socioeconomic status, white matter, and executive function in children. Brain Behav. 6, e00531. https://doi.org/10.1002/brb3.531

Van Beek, L., Ghesquière, P., Lagae, L., De Smedt, B., 2014. Left fronto-parietal white matter correlates with individual differences in children's ability to solve additions and multiplications: a tractography study. NeuroImage 90, 117–127. https://doi.org/10.1016/j.neuroimage.2013.12.030

van Buuren, S., 2007. Multiple imputation of discrete and continuous data by fully conditional specification. Stat. Methods Med. Res. 16, 219–242. https://doi.org/10.1177/0962280206074463

van Buuren, S., Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. J. Stat. Softw. 45. https://doi.org/10.18637/jss.v045.i03

van de Ven, V.G., Formisano, E., Prvulovic, D., Roeder, C.H., Linden, D.E.J., 2004. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum. Brain Mapp. 22, 165–178. https://doi.org/10.1002/hbm.20022

van den Heuvel, M.P., Hulshoff Pol, H.E., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur. Neuropsychopharmacol. J. Eur. Coll. Neuropsychopharmacol. 20, 519–534. https://doi.org/10.1016/j.euroneuro.2010.03.008

Van Dijk, K.R.A., Hedden, T., Venkataraman, A., Evans, K.C., Lazar, S.W., Buckner, R.L., 2010. Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. J. Neurophysiol. 103, 297–321. https://doi.org/10.1152/jn.00783.2009

Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880. https://doi.org/10.1109/10.623056

Vandermosten, M., Boets, B., Wouters, J., Ghesquière, P., 2012. A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. Neurosci. Biobehav. Rev. 36, 1532–1552. https://doi.org/10.1016/j.neubiorev.2012.04.002

Vandewater, E.A., Bickham, D.S., Lee, J.H., Cummings, H.M., Wartella, E.A., Rideout, V.J., 2005. When the television is always on: Heavy television exposure and Young children's development. Am. Behav. Sci. 48, 562–577. https://doi.org/10.1177/0002764204271496

Varoquaux, G., Craddock, R.C., 2013. Learning and comparing functional connectomes across subjects. NeuroImage, Mapping the Connectome 80, 405–415. https://doi.org/10.1016/j.neuroimage.2013.04.007

Verhoeven, L., Leeuwe, J. van, 2008. Prediction of the development of reading comprehension: a longitudinal study. Appl. Cogn. Psychol. 22, 407–423. https://doi.org/10.1002/acp.1414

Verstynen, T.D., Weinstein, A.M., Schneider, W.W., Jakicic, J.M., Rofey, D.L., Erickson, K.I., 2012. Increased Body Mass Index Is Associated With a Global and Distributed Decrease in White Matter Microstructural Integrity. Psychosom. Med. 74, 682–690. https://doi.org/10.1097/PSY.0b013e318261909c

Vértes, P.E., Bullmore, E.T., 2015. Annual Research Review: Growth connectomics – the organization and reorganization of brain networks during normal and abnormal development. J. Child Psychol. Psychiatry 56, 299–320. https://doi.org/10.1111/jcpp.12365

Vestergaard, M., Madsen, K.S., Baaré, W.F.C., Skimminge, A., Ejersbo, L.R., Ramsøy, T.Z., Gerlach, C., Akeson, P., Paulson, O.B., Jernigan, T.L., 2011. White matter microstructure in superior longitudinal fasciculus associated with spatial

working memory performance in children. J. Cogn. Neurosci. 23, 2135–2146. https://doi.org/10.1162/jocn.2010.21592

Vissers, M.E., Cohen, M.X., Geurts, H.M., 2012. Brain connectivity and high functioning autism: a promising path of research that needs refined models, methodological convergence, and stronger behavioral links. Neurosci. Biobehav. Rev. 36, 604–625. https://doi.org/10.1016/j.neubiorev.2011.09.003

Voillet, V., Besse, P., Liaubet, L., San Cristobal, M., González, I., 2016. Handling missing rows in multi-omics data integration: multiple imputation in multiple factor analysis framework. BMC Bioinformatics 17. https://doi.org/10.1186/s12859-016-1273-5

Voss, M.W., Prakash, R.S., Erickson, K.I., Basak, C., Chaddock, L., Kim, J.S., Alves, H., Heo, S., Szabo, A.N., White, S.M., Wójcicki, T.R., Mailey, E.L., Gothe, N., Olson, E.A., McAuley, E., Kramer, A.F., 2010. Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. Front. Aging Neurosci. 2. https://doi.org/10.3389/fnagi.2010.00032

Waller, R., Dotterer, H.L., Murray, L., Maxwell, A.M., Hyde, L.W., 2017. White-matter tract abnormalities and antisocial behavior: A systematic review of diffusion tensor imaging studies across development. NeuroImage Clin. 14, 201–215. https://doi.org/10.1016/j.nicl.2017.01.014

Wang, X., Han, Z., He, Y., Liu, L., Bi, Y., 2012. Resting-state functional connectivity patterns predict Chinese word reading competency. PloS One 7, e44848. https://doi.org/10.1371/journal.pone.0044848

Weatherholt, T.N., Harris, R.C., Burns, B.M., Clement, C., 2006. Analysis of attention and analogical reasoning in children of poverty. J. Appl. Dev. Psychol. 27, 125–135. https://doi.org/10.1016/j.appdev.2005.12.010

Webber, M.P., Huxley, P.J., 2007. Measuring access to social capital: the validity and reliability of the Resource Generator-UK and its association with common mental disorder. Soc. Sci. Med. 1982 65, 481–492. https://doi.org/10.1016/j.socscimed.2007.03.030

Wechsler, D., 2011. Wechsler abbreviated scale of intelligence- second edition. Pearson, Bloomington, MN.

Wellcome Trust Centre for Neuroimaging, 2014. SPM12 - Statistical Parametric Mapping [WWW Document]. SPM. URL https://www.fil.ion.ucl.ac.uk/spm/software/spm12/ (accessed 9.24.18).

Welsh, J.A., Nix, R.L., Blair, C., Bierman, K.L., Nelson, K.E., 2010. The Development of Cognitive Skills and Gains in Academic School Readiness for Children from Low-Income Families. J. Educ. Psychol. 102, 43–53. https://doi.org/10.1037/a0016738

Werner, E.E., 1993. Risk, resilience, and recovery: Perspectives from the Kauai Longitudinal Study. Dev. Psychopathol. 5, 503–515. https://doi.org/10.1017/S095457940000612X

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., Bryant, D., 2011. Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. Dev. Neuropsychol. 36, 162–180. https://doi.org/10.1080/87565641.2010.549980

Wilson, P., Bradshaw, P., Tipping, S., Henderson, M., Der, G., Minnis, H., 2012. What predicts persistent early conduct problems? Evidence from the Growing Up in Scotland cohort. J Epidemiol Community Health jech-2011-200856. https://doi.org/10.1136/jech-2011-200856

Witten, D.M., Tibshirani, R.J., 2009. Extensions of Sparse Canonical Correlation Analysis with Applications to Genomic Data. Stat. Appl. Genet. Mol. Biol. 8. https://doi.org/10.2202/1544-6115.1470

Wold, H., 1982. Soft modeling : the basic design and some extensions. Syst. Indirect Obs. Causality Struct. Predict., Systems under indirect observation : causality, structure, prediction ; 2 2.

Wold, H., 1975. Path Models with latent variables: The NIPALS approach. Acad. Press.

Wood, A.M., White, I.R., Royston, P., 2008. How should variable selection be performed with multiply imputed data? Stat. Med. 27, 3227–3246. https://doi.org/10.1002/sim.3177

Woodcock, R.W., McGrew, K.S., Mather, N., 2001. Woodcock-Johnson III tests of achievement. Riverside Publishing., Itasca, IL.

Woolrich, M., Hunt, L., Groves, A., Barnes, G., 2011. MEG beamforming using Bayesian PCA for adaptive data covariance matrix regularization. NeuroImage 57, 1466–1479. https://doi.org/10.1016/j.neuroimage.2011.04.041

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M., 2009. Bayesian analysis of neuroimaging data in FSL. NeuroImage 45, S173-186. https://doi.org/10.1016/j.neuroimage.2008.10.055

Yanagisawa, K., Masui, K., Furutani, K., Nomura, M., Yoshida, H., Ura, M., 2012. Family socioeconomic status modulates the coping-related neural response of offspring. Soc. Cogn. Affect. Neurosci. nss039. https://doi.org/10.1093/scan/nss039

Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. J. R. Stat. Soc. Ser. B 68, 49–67.

Zajonc, R.B., Markus, G.B., 1975. Birth order and intellectual development. Psychol. Rev. 82, 74–88. https://doi.org/10.1037/h0076229

Zalesky, A., Fornito, A., Bullmore, E.T., 2010. Network-based statistic: Identifying differences in brain networks. NeuroImage 53, 1197–1207. https://doi.org/10.1016/j.neuroimage.2010.06.041

Zhang, M., Li, J., Chen, Chuansheng, Xue, G., Lu, Z., Mei, L., Xue, H., Xue, F., He, Q., Chen, Chunhui, Wei, M., Dong, Q., 2014. Resting-state functional connectivity and reading abilities in first and second languages. NeuroImage 84. https://doi.org/10.1016/j.neuroimage.2013.09.006

Zhou, W., Xia, Z., Bi, Y., Shu, H., 2015. Altered connectivity of the dorsal and ventral visual regions in dyslexic children: a resting-state fMRI study. Front. Hum. Neurosci. 9. https://doi.org/10.3389/fnhum.2015.00495

Ziegler, G., Dahnke, R., Winkler, A.D., Gaser, C., 2013. Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. NeuroImage 82, 284–294. https://doi.org/10.1016/j.neuroimage.2013.05.088

10 APPENDICES

APPENDIX A: THE ENVIRONMENT QUESTIONNAIRES

Primary caregivers completed a number of questionnaires, created to collect information on the environment domains. The following questionnaire was the main questionnaire completed during unit testing:

Subject Number:

Parental Questionnaire

Please read the following before filing in this questionnaire:

This questionnaire is designed to be completed by the child's primary caregiver (the adult that spends the most time with the child/caring for the child. This person should be someone that is living with the child). If you are not the child's primary caregiver please only answer the questions that you feel comfortable answering on behalf of the primary caregiver. Please then notify the individual that gave you this questionnaire, so that they can ensure the remaining questions are answered by the primary caregiver.

The aim of this questionnaire is to gather information about your home life. There are no right or wrong answers. Your responses are also strictly confidential, so please be honest. Please try and answer all the questions. If you would rather someone complete the questionnaire with you, or if you have any questions, then please ask.

Section 1: Who is completing this questionnaire?
1) What is your relationship to the child?
2) Are you the child's primary caregiver (the person that spends the most time looking after the child)? □ Yes   □ No If no, please specify who the child's primary caregiver is (e.g. the child's mother, the child's grandfather etc.) and why they were unable to bring their child to the unit today.

Section 2: General
1) What is your child's Nationality?
2) What is your child's first spoken language (the language they use most often)?
3) Do you (the primary caregiver) speak any other languages at home?   □ Yes   □ No If yes, please specify which languages
4) Does your child speak any other languages at home?   □ Yes   □ No If yes, please specify which languages

5) Please select the option below that you feel best applies to the communication your child takes part in at home:

□ Speaks English all of the time
□ Speaks English the majority of the time
□ Speaks English approximately half of the time
□ Rarely speaks English
□ Never speaks English

6) Which of these do you most identify as?

□ Christian   □ Muslim   □ Jewish   □ Hindu
□ Buddhist   □ Sikh   □ Atheist (no religion)   □ Agnostic (undecided about religion)
□ Other (please specify)

7) Do you consider yourself to be actively participating in your religion, if any (i.e. attending meetings or practicing religious worship)

□ I am extremely active   □ A am somewhat active   □ I am occasionally active   □ I am not at all active
□ I do not actively participating in any religion because I am atheist or agnostic.

8) Would you consider your child to hold the same religious beliefs as you? If not please state why?

9) What would you consider to be your child's best qualities? Please elaborate as far as possible.

Section 2: Home And Community Environment

1) Does your child have a mobile phone?   □ **Yes**   □ **No**

2) Please select which of the following most applies to you:

□ We have no computers/laptops/netbooks/tablets in the home
□ We have one computer/laptop/netbook/tablet in the home
□ We have more than one computer/laptop/netbook/tablet in the home
□ Each individual person has a computer/laptop/netbook/tablet in the home

3) Do you have internet in your household?

□ **Yes**   □ **No**

If yes, how many hours a day does your child use the internet on average?
□ Less than 1 hour
□ 1 hour
□ 2 hours
□ 3 hours
□ 4 hours +
□ Don't know

Which of the following activities does your child use the internet for? Tick all that apply
□ Playing games
□ Social media (E.g. Facebook)
□ Learning new things (E.g. Wikipedia or tutorials on Youtube)
□ Homework
□ Watching TV and films
□ Communicating with others (E.g. email or Skype)
□ Other use. Please specify:

4) Does your child have free access to a computer/laptop/netbook/tablet?	□ **Yes**	□ **No**
If no, do they have monitored or supervised access?	□ **Yes**	□ **No**

5) Do you have a television in your house?

□ Yes, one television   → Go to question 6)
□ Yes, two televisions   → Go to question 6)
□ Yes, more than two televisions   → Go to question 6)
□ No. We don't own a television   → Go to question 10)

6) Does your child have a television or computer in their bedroom?	□ **Yes**	□ **No**

7) Do you have any rules about what your child can or can't watch on the television or how long they are allowed to watch for?

□ **Yes**   □ **No**

8) How many hours each day on average does your child watch the television?

□ Less than 1 hour
□ 1 hour
□ 2 hours
□ 3 hours
□ 4 hours +
□ Don't know

9) Do you have any games consoles in the house that your child has access to play with?

□ **Yes**   □ **No**

If yes, do you have any rules about what your child can or can't play or how long they are allowed to play for?

□ **Yes**   □ **No**

10) Please select which of the following most applies to you?

□ We have no books in the home
□ We have a few books in the home
□ We have a number of books in the home, but none suited to children
□ We have a number of books in the home, some suited to children
□ We have a large number of books in the home, covering a range of subjects and reading abilities

11) Which of these best describes your situation?

□ I actively encourage my child to read
□ My child has access to books at home and can read when they choose to
□ My child reads at school but not at home
□ I don't think reading is important in the home environment

12) Which of the following most applies to you?

□ Myself or another family member reads out loud with my child every day
□ Myself or another family member reads out loud with my child most days
□ Myself or another family member reads out loud with my child occasionally
□ We do not read out loud at home

13) On average, how many minutes is your child read to at home per day?

□ Less than 5 minutes/day
□ 5 – 10 minutes/day
□ 11 – 20 minutes/day
□ 21 – 40 minutes/day
□ More than 40 minutes/day

On average, how many minutes does your child read alone at home per day?

□ Less than 5 minutes/day
□ 5 – 10 minutes/day
□ 11 – 20 minutes/day
□ 21 – 40 minutes/day
□ More than 40 minutes/day

14) Which of the following most applies to you?

□ I really enjoy reading and often read on my own for fun when I get the chance
□ I quite enjoy reading and sometimes read on my own for fun
□ I neither like or dislike reading on my own
□ I don't really enjoy reading and rarely read on my own for fun
□ I dislike reading and never read on my own for fun

15) Which of the following most applies to your child?

□ They really enjoy reading and often read for fun
□ They quite enjoy reading and sometimes read for fun
□ They neither like or dislike reading
□ They don't really enjoy reading and rarely read for fun
□ They dislike reading and never read for fun

16) How often does your child use the local library?

□ More than once a week
□ Once a week
□ Once a fortnight
□ Once a month
□ Once every 3months
□ Once every 6 months
□ Once a year
□ Less than once a year
□ Never

17) Does someone in the family get a daily newspaper?	□ **Yes**	□ **No**

18) Do you ever do any of the following with your child?

□ Watch TV together
□ Play with puzzles or board games together
□ Draw together
□ Sing together
□ Play inside together
□ Play outside together
□ Cook together
□ Discuss TV programs or books together
□ Watch sporting events together
□ Go to musical or drama performance together
□ Visit a museum together

19) What would your ideal family day out be?

20) Does your child have a designated bedroom?	□ **Yes**	□ **No**

If yes, does your child share this bedroom with another person? □ **Yes**   □ **No**

If yes, is the person sharing the room with the child more than five years older or younger than the child?

□ **Yes**   □ **No**

21) Does your child have a bicycle?	□ **Yes**	□ **No**
22) Does your child have access to a garden/green space?	□ **Yes**	□ **No**

23) On average, how many days a week does your child spend some time playing outside?

24) Does your child regularly play with other children outside of school?	□ **Yes**	□ **No**

25) Is your child allowed to go out unsupervised to play with friends?	□ **Yes**	□ **No**

If yes, do you know where your child is going?

□ Always □ Most of the time □ About half the time □ Occasionally □ Rarely

26) What is the postcode for the house that the child lives in? (Providing us with this helps us to access local information e.g. proximity to children's play areas/green space)

27) Does your child have a set dinner time?	□ **Yes**	□ **No**

28) Does your child sit at a table to eat dinner?	□ **Yes**	□ **No**

29) Approximately how many days a week do you eat dinner as a family?

□ 7 days per week
□ 6 days per week
□ 5 days per week
□ 4 days per week
□ 3 days per week
□ 2 days per week
□ 1 day per week
□ No days per week

30) How well do you think you listen to this child?

□Very well □ Fairly well □ I don't know □ Fairly badly □ Very badly

31) How often in the last six months have you done any of the following?

Discussed with your child what they have done wrong and why you are disappointed
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

Shouted at your child
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

Grounded your child/prevented your child going out to play with friends
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

Hit/smacked your child
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

Taken away TV time, computer time, child's phone etc.
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

Given your child a time out/sent them to their room
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

Let them get away with causing trouble because they don't respond to disciplining
□ Never □ Very Rarely □ Occasionally □ Frequently □ Very often

32) Some parents spend time teaching their children new skills while other parents believe children learn best on their own. Which most closely describes your attitude?

□ Children need to be taught new skills
□ Children learn best on their own

33) Do you worry about how your child performs at school?

□ Never   □ Rarely   □ Sometimes   □ Often   □ All the time

34) Is your child expected to make his/her bed?   □ **Yes**   □ **No**

35) Is your child expected to clean his/her room?   □ **Yes**   □ **No**

36) Is your child expected to bathe him/herself?   □ **Yes**   □ **No**

37) Is your child expected to do routine chores such as mowing the lawn, helping to make dinner or washing the dishes?

□ **Yes**   □ **No**

38) Please indicate how much you agree with the following statements

a) I feel like I belong to the neighbourhood I live in.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

b) The friendships and associations I have with other people in my neighbourhood mean a lot to me.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

c) If I needed advice about something I could go to someone in my neighbourhood.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

d) I borrow things and exchange favours with my neighbours.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

e) I would be willing to work together with others on something to improve my neighbourhood.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

f) I like to think of myself as similar to the people who live in this neighbourhood.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

g) I regularly stop and talk with people in my neighbourhood.

□ Strongly disagree   □ Disagree   □ Neither agree or disagree   □ Agree   □ Strongly agree

39) How many rooms are there in your home?

40) How many bedrooms are there in your home?

Section 3: Child Health and Education

1) In general how would you say your child's current health is?

□ Excellent   □ Very Good   □ Good   □ Fair   □ Poor

2) In general how would you say your child's health throughout their life has been?

□ Excellent   □ Very Good   □ Good   □ Fair   □ Poor

3) On average, how many hours sleep do you think your child gets a night?

4) Was the child a twin, triplet or more?

□ Single birth
□ Twin
□ Triplet
□ Other:

5) Whilst pregnant did the mother of the child do any of the following?

□ Consume alcohol
□ Take prescription medication
□ Take non-prescription medications
□ Take illegal drugs
□ Smoke
□ None of the above

6) Which of the following best describes your child's approach to school:

□ My child has always disliked school
□ My child has always liked school
□ My child's attitude to school has been dependant on his/her teacher
□ Over the years, my child seems to look forward to school more
□ Over the years, my child seems to look forward to school less
□ My child has always found school to be okay

7) How often on average do you talk to your child about his/her day or activities at school?

□ Everyday   □ Most days   □ Some days   □ Rarely   □ Never

---

8) If your child has done particularly well on a school project, which of the following are you most likely to say?

□ Well done, you're so clever!
□ Well done, you worked so hard on that!
□ Nothing at all

---

9) How often do you or your partner help your child with their homework?

□ Everyday   □ Most days   □ Some days   □ Rarely   □ Never

---

10) Do you have a designated space in your home for your child to do their homework?   □ **Yes**   □ **No**

---

11) Which of the following best describes your opinion?

□ It is important that my child does their homework. I know when they have homework and help them to complete it daily. We have dedicated homework time.
□ It is important that my child does their homework. I prompt my child to do it regularly but they do it independently in their own time.
□ My child is responsible for completing their own homework. I do not often prompt them to do it.
□ To my knowledge my child does not get set homework.
□ I do not think work is important in the home environment so my child does not do homework at home.

---

12) Does your child receive any educational tutoring outside of school?   □ **Yes**   □ **No**

If yes, for how many hours a week?

---

13) Since your child started school, in an average year how many days does your child not attend school on a normal school day?

□ My child has less than 5 days off school on average each year
□ My child has 6 – 14 days off school on average each year
□ My child has 14 – 21 days off school on average each year
□ My child has > 21 days off school on average each year

---

14) Does your child play a musical instrument?   □ **Yes**   □ **No**

---

15) Does your child take part in any extra-curricular activities provided by their school or outside of school?

□ **Yes**   □ **No**

If yes, please specify (e.g. football, art club etc.)

---

16) Since your child started school have you ever attended a Parent's Evening/Parents-Teacher Meeting?

□ **Yes**   □ **No**

17) Does your child eat breakfast every day?

□ **Yes**　　　　□ **No**　　　　□ **Don't know**

18) Does your child eat dinner every night?

□ **Yes**　　　　□ **No**

19) On average how many portions of fruit and vegetables does your child eat each day?

□ None
□ 1 or 2
□ 3 or 4
□ 5 or more

20) On average how many of each of the following does your child eat per day?

a) packets of crisps
□ None
□ 1 or 2
□ 3 or 4
□ 5 or more

b) chocolate bars
□ None
□ 1 or 2
□ 3 or 4
□ 5 or more

c) handful of sweets
□ None
□ 1 or 2
□ 3 or 4
□ 5 or more

d) biscuits/cakes
□ None
□ 1 or 2
□ 3 or 4
□ 5 or more

21) Has your first child ever been home schooled?	□ Yes	□ No
Has your second child ever been home schooled?	□ Yes	□ No

22) Which school is your first child at?

Which school is your second child at?

Section 4: Additional Caregiver Information

1) Which best describes your current occupation status?

□ Full time employee → Go to question 4)

□ Part time employee → Go to question 4)

□ Part-time employee and part-time student → Go to question 4)

□ Self-employed → Go to question 4)

□ Employed on a zero hour contract → Go to question 4)

□ Full time student → Go to question 3)

□ Part-time student without job → Go to question 2)

□ Part-time student with job → Go to question 4)

□ Out of work → Go to question 2)

---

2) What is the main reason you are currently out of paid work?

□ Myself or my partner have had another baby

□ I have a child that doesn't go to school. I prefer to stay home and look after them

□ I have a child that doesn't go to school and I can't get suitable childcare

□ It's not worthwhile because of the cost of childcare for my children whilst I am at work

□ I have other caring responsibilities □ My own ill health

□ My partner earns enough to support us

□ I don't have the right kind of education/training to get a job

□ There are no suitable jobs available

□ I can't find a job that interests me

□ I can't find a job with enough flexibility

□ Transport problems

□ I'd lose government benefits if I worked

□ Other reason. Please specify

---

3) How long have you currently been out of paid work?

---

4) Excluding any period of maternity/paternity leave which of the following most applies to you:

□ I have not worked since the birth of my child

□ I have worked a little since the birth of my child but not for more than a one year period

□ I have been in work for more than a one year period

□ I have been in work for at least three years since the birth of my child

□ I have been in work for at least five years since the birth of my child

□ I have been in work almost all of the time and have only recently stopped working

---

5) On average, how many hours do you work each week? Please complete for your last main job if you are currently out of work

6) Do you ever work after 6pm at night or at weekends? Please complete for your last main job if you are currently out of work

□ **Yes**   □ **No**

If yes please select which of the following most applies to you:

□ Permanent night shift
□ Work after 6 or at the weekends 4-6 days a week
□ Work after 6 or at the weekends 1-3 days a week
□ Work after 6 or at the weekends once a month
□ Work after 6 or at the weekends less than once a month

7) If you are living with a partner, do they ever work after 6pm at night or at weekends?

□ Yes   □ No   □ N/A

If yes please select which of the following most applies to them:

□ Permanent night shift
□ Works after 6 or at the weekends 4-6 days a week
□ Works after 6 or at the weekends 1-3 days a week
□ Works after 6 or at the weekends once a month
□ Works after 6 or at the weekends less than once a month

8) Does your household receive any of the following benefits? (Please tick all that apply)

□ Housing Benefit
□ Child Benefit
□ Council Tax Reduction
□ Guardian's Allowance
□ Jobseeker's Allowance (JSA)
□ Employment and Support Allowance (ESA)
□ Income Support
□ Disability Living Allowance (DLA) or Personal Independence payment (PIP)
□ Healthy Start
□ Parents Learning Allowance
□ Childcare and child tax credits
□ Carers credit or Carers allowance
□ Winter fuel payment
□ Reduced Earnings Allowance
□ Other, please state:

9) Does your child receive free school meals?

□ Yes   □ No

10) Does the child have another caregiver, parental figure or parent **not** living in the child's home?

□ **Yes**   → Go to question 11
□ **No**   → Go to question 16

11) Which of the following best describes the caregiver or parental figure that lives outside of the home?

□ a biological parent
□ A step parent
□ A partner of the child's primary caregiver
□ A friend of the child's primary caregiver
□ Other, please specify:

12) What is the current occupation of the main caregiver or parental figure that lives outside of the home?

13) Does this individual support the child in any of the following ways? (Please tick all that apply)

□ Financial support
□ Help with school work
□ Advice and emotional support for the child
□ Plays a significant role in their upbringing
□ None of the above

14) Is the additional caregiver or parental figure that lives outside of the home able to read and write?

□ **Yes**   □ **No**

15) On average, how many days each week does the child visit this caregiver/parental figure?

16) Who provides regular care for your child when your child is not at school? Please tick all the types of care that your child receives.

Care by:
□ Yourself
□ A family member or spouse/partner who lives with you
□ Your spouse/partner who lives elsewhere
□ Parent of the child who lives elsewhere
□ Other relative 16 years or over (including siblings)
□ Other person 16 years or over e.g. nanny, au pair, friend, neighbour
□ Relative under 16 years (including siblings)
□ Other person under 16 years
□ Child cares for self
□ Child care centre or a before or after-school club
        Please give the name of the child care centre if possible:

□ Other. Please specify:

17) How old was your child when you first started using any type of regular childcare for him/her?

□ Less than 6 months
□ 7 – 12 months
□ 13 – 18 months
□ 19 months – 3 years
□ Over 3 years – 5 years
□ Over 5 years
□ I have never used regular child care

18) How often does your child get to see the following people:

Grandparents	**Aunts and uncles**	**Cousins**	**Family friends**
□ Often	□ Often	□ Often	□ Often
□ Occasionally	□ Occasionally	□ Occasionally	□ Occasionally
□ Rarely	□ Rarely	□ Rarely	□ Rarely
□ Never	□ Never	□ Never	□ Never

19) Would you consider yourself to have a good relationship with your family?

□Very good   □ Fairly good   □ I don't know   □ Fairly bad   □ Very bad

20) Would you consider yourself to have a good relationship with your child?

□Very good   □ Fairly good   □ I don't know   □ Fairly bad   □ Very bad

21) Overall, how do you feel about the amount of support or help you get from family or friends living elsewhere?

□ I get enough help
□ I don't get enough help
□ I don't get any help at all
□ I don't need any help

22) How often do you feel you need support or help but can't get it from anyone?

□ Very often   □ Often   □ Sometimes   □ Never   □ I don't need it

23) How often do you have a drink containing alcohol?

□ Everyday
□ 4 - 6 times a week
□ 2 - 3 times a week
□ Once a week
□ 2 - 3 times a month
□ Monthly or less
□ I don't drink alcoholic drinks

24) How many units do you have on a standard day when drinking, using the following estimates?
(1 pint of lager/cider = 3 units)
(medium glass of wine = 2 units)
(single shot measure = 1 unit)

25) How often do you experience conflict with your partner in a normal week?

□ Every day   □ Some days   □ Rarely   □ Not at all

26) On average, would you say that you find life stressful?   □ **Yes**   □ **No**

27) Since the birth of your child have you experienced any of the following? Please tick all that apply as this will allow us to get a standardised measure of stressful events:

□ Death of spouse
□ Divorce
□ Marital separation
□ Jail term
□ Death of close family member
□ Personal injury or illness
□ Marriage
□ Fired at work
□ Marital reconciliation
□ Retirement
□ Change in health of family member
□ Pregnancy
□ Sex difficulties
□ Gain of new family member
□ Business readjustment
□ Change in financial state
□ Death of close friend
□ Change to different line of work
□ Change in number of arguments with spouse
□ Take out a mortgage over £60,000
□ Unable to pay mortgage or loan
□ Change in responsibilities at work
□ Son or daughter leaving home
□ Trouble with in-laws
□ Outstanding personal achievement
□ Partner beginning or stopping work
□ Beginning or ending education

□ Change in living conditions
□ Revision of personal habits
□ Trouble with boss
□ Change in work hours or conditions
□ Change in residence
□ Change in school
□ Change in recreation
□ Change in religious activities
□ Change in social activities
□ Take out a mortgage or loan less than £60,000
□ Change in sleeping habits
□ Change in number of family get-togethers
□ Change in eating habits
□ Vacation
□ Observed a major festival or celebration e.g. Christmas, Diwali, Thanksgiving etc.
□ Minor violations of the law
□ Suffered a serious injury, illness or assault
□ Returned to work after a period of absence
□ Significantly increased work hours
□ Significantly decreased work hours
□ Sought work unsuccessfully for more than 1 month
□ Been away from home a lot
□ Had an alcohol or drug problem
□ Had violent arguments
□ Been burgled

Which of the events selected did you experience in the last 12 months?

28) Would you consider yourself to have a disability? □ **Yes** □ **No**

If, yes would you consider yourself to have

□ a physical disability
□ an emotional or behavioural disability
□ Other. Please specify:

29) We'd like to know how important various things are to your sense of who you are.

Please think about each of the following and tick the box that indicates whether you think it is very important, fairly important, not very important or not at all important to your sense of who you are.

a) Your profession?

□ Very important □ Fairly important □ Not very important □ Not at all important

b) Your level of education?

□ Very important □ Fairly important □ Not very important □ Not at all important

c) Your family

□ Very important □ Fairly important □ Not very important □ Not at all important

30) Think of this ladder as representing where the people stand in terms of status in the United Kingdom.

At the TOP of the ladder are the people who are the best off—those who have the most money, the most education, and the most respected jobs.

At the BOTTOM are the people who are the worst off—who have the least money, least education, and the least respected jobs or no job.

The higher up you are on this ladder, the closer you are to the people at the very top. The lower you are, the closer you are to the people at the very bottom.

Where would you place yourself on this ladder, compared to all the other people in the United Kingdom?

Please place a large "X" directly on the rung where you think you stand.

31) If you could offer one piece of advice to new parents, what would it be?

Thank you for taking the time to fill in this questionnaire
The answers you have given will remain strictly confidential

The following questionnaires were sent to participants before attending the unit to complete at home as they involve items that participants might not know without looking up. Sending these home helped to ensure that we collected as accurate information as possible.

Subject Number:

MRC | Cognition and Brain Sciences Unit

A.C.E
Attention Cognition Education

To be completed at home: Parental Questionnaire

Please read the following before filling in this questionnaire:

This questionnaire should be completed by the child's primary caregiver (the adult that spends the most time with the child/caring for the child). The aim of this questionnaire is to gather information about your home life. We will ask further questions when you come in to the unit, but this questionnaire contains questions that you may want to look up whilst at home. There are no right or wrong answers. Your responses are also strictly confidential, so please be honest and please try and answer all the questions.

Please bring the completed questionnaire with you when you visit the CBU for your child's MEG session. If you have any difficulties at all completing this questionnaire, please still bring it along with you and we can help you to complete it when you are here.

Please specify your relationship to the study child in the space provided below. You should only complete this questionnaire if you are the child's primary caregiver (the adult that spends the most time with the child/caring for the child).

Section 1: Caregiver/Parent Information

1) How much did your child weigh at birth?

2) Was your child born pre-term (before their due date)?

☐ My child was born before the 32nd week of pregnancy
☐ My child was born between the 32nd and 37th week of pregnancy
☐ My child was born between the 37th and 42nd week of pregnancy
☐ My child was born after the 42nd week of pregnancy

3) How old was the mother when the child was born?

4) How old was the father when the child was born?

5) How old was your child when the mother went back to paid work after giving birth?

☐ Less than 6 months
☐ 6 months to 12 months
☐ 13 months to 18 months
☐ 19 months to 3 years
☐ 3 years to 5 years
☐ Older than 5 years
☐ The mother has not returned to work
☐ The mother never worked
☐ Don't know

Subject number:

MRC | Cognition and Brain Sciences Unit

A.C.E
Attention Cognition Education

Household Dynamics

Please ask each individual living in the same household as your child to complete one of these questionnaires.

These questionnaires are designed to gather information about the individuals that your child lives with. If an individual living in the household is unavailable to complete the questionnaire or is under the age of 16, we ask that the adult that knows most about them complete this questionnaire on their behalf. All responses are strictly confidential, so please be honest and please try to answer all the questions. If you have any difficulties completing this questionnaire, please bring it along to the unit and we can help you complete this.

Section 1: Family dynamics
1) First name of individual
2) Gender: ☐ Male ☐ Female
3) Date of Birth: (dd/mm/yyyy)
4) Are you temporarily living away from home? (for more than one month) ☐ Yes ☐No If yes, why? ☐ Work reasons ☐ Hospital/Illness ☐ Boarding school/college/university ☐ Other reason, please specify below:
5) What is your relation to the primary caregiver/are you the child's primary caregiver? If you did not identify yourself as the primary caregiver, would you consider yourself to be the child's secondary caregiver? ☐ Yes ☐No

6) Are you able to read and write?

☐ Yes   ☐No

7) What is the highest level of education you have **completed**?

☐ Primary school
☐ Some secondary school *
☐ CSEs/O levels/GCSEs/Level 1 or 2 NVQ
☐ AS levels/ Level 3 NVQ
☐ A levels/ Level 4 NVQ/ Level 3 BTEC/International Baccalaureate
☐ Apprenticeship
☐ Certificate of higher education/diploma/foundation degree
☐ Bachelor's degree/ Professional certificate
☐ Master's Degree/PGCE
☐ PhD/Doctorate
☐ Other, please specify:
☐ None of the above

* If you ticked some secondary school, please specify how old you were when they left/what year you are currently in:

8) Which of the following best describes your marital status?

☐ Single   ☐ Married or Civil Partnership   ☐ Living with partner
☐ Divorced   ☐ Widowed   ☐ Separated

Section 2: Occupation

1) Are you currently in paid work?

☐ Yes
Please answer the following questions for your current job
☐ No, but I have been in paid work within the past year:
 please answer the following questions for your last main job
☐ No, I have been unemployed for the past year:
 please answer the following questions for your last main job
☐ No, I am a full time student:
if you have previously been in paid work, please answer the following for your last main job. If you have not previously been in paid work, please do not answer the remaining questions
☐ No, I have always been unemployed:
please do not answer the remaining questions

2) What does the firm/organisation you work/worked for mainly make or do? Please be as descriptive as possible.

3) What is/was your main job?

4) What did/do you mainly do in your job? Please be as descriptive as possible.

5) Did you need any special qualifications or training to do this job? Please give the names of qualifications if known.

6) Are you/were you working as an employee or are you/were you self-employed?

☐ Employee: Please complete questions 7 and 8
☐ Self-employed: Please complete questions 9 and 10

7) In your job, do you/did you have any formal responsibility for supervising the work of other employees?

8) How many people work/worked for your employer at the place where you work/worked? Please give the number of total employees working on that site and not just the number in your department etc.
☐ 1 - 9
☐ 10 - 24
☐ 25 - 499
☐ 500 or more

Please do not complete questions 9 and 10 if you completed questions 7 and 8

9) Are you/were you working on your own or did you have employees?

☐ Working/worked on your own or with a partner or partners
☐ Working/worked with employees

10) How many people did/do you employ at the place where you work/worked?

☐ 1 - 9 people
☐ 10 - 24 people
☐ 25 - 499 people
☐ 500 or more people

Thank you for taking the time to fill in this questionnaire.

The answers you have given will remain strictly confidential.

Finally, the following questionnaire was completed by the child participants:

Subject Number:

Child Questionnaire

These questions are about how you feel about different aspects of your life. Circle the face that best describes how you feel about the different things listed. If you feel completely happy with something, circle the happy face by number 1. If you feel very unhappy about something, circle the sad face by number 7. There are no right or wrong answers.

A) Your school work?

Very happy   Very sad

1   2   3   4   5   6   7

B) Your appearance?

1   2   3   4   5   6   7

C) Your parents?

1   2   3   4   5   6   7

D) Your brothers or sisters?

1   2   3   4   5   6   7

E) Your friends?

1   2   3   4   5   6   7

F) The school you go to?

1   2   3   4   5   6   7

G) Which best describes how you feel about your life as a whole?

1   2   3   4   5   6   7

APPENDIX B: THE ENVIRONMENT DOMAINS

Table 8 Variables removed from analysis as over 90% of subjects had the same value for these variables

Care by other nonrelative 16years or over
Care by other nonrelative under 16 years
Care by other relative under 16years
Care by partner
Care for themselves
Caregiver has emotional or behavioural disability
Caregiver has other disability
Caregiver has physical disability
Ever been home schooled
Extra tutoring outside school
Has internet
Importance of family to caregiver
Mother consumed alcohol in pregnancy
Mother took illegal drugs in pregnancy
Mother took nonprescription med in pregnancy
Mother took prescription med in pregnancy
Parental beliefs of child learning
Room shared person more 5yr age gap
Twin or triplet

Table 9 The environmental domains and their corresponding questionnaire items. Skewed variables listed in Table 8 have been removed. The final column indicates whether each item was removed from the factor analyses to create the final set of domains.

Domain	Questionnaire Item	Type of data	Included?
Additional info	Gender	Binary	NA
Additional info	Handedness	Binary	NA
Additional info	Age in years	Continuous	NA
Income	Equivalised income	Continuous	Y
Parent education	Parent education av	Continuous	Y
Parent occupation	Parent occupation av	Continuous	Y
Subjective SES	Subjective SES	Continuous	Y
Number adults	Number adults	Continuous	Y
Number siblings	Number siblings	Continuous	Y
Primary work hours	Primary work hours	Continuous	Y
RG personal skills	RG personal	Continuous	Y
Acquaintance skills	RG domestic	Continuous	Y
Acquaintance skills	RG expert	Continuous	Y
Acquaintance skills	RG skills	Continuous	Y
Acquaintance skills	RG problems	Continuous	Y
Neighbourhood SES	Income decile	Continuous	Y
Neighbourhood SES	Employment decile	Continuous	Y
Neighbourhood SES	Education and Skills decile	Continuous	Y
Neighbourhood SES	Health and disability decile	Continuous	Y
Neighbourhood SES	Crime decile	Continuous	Y

Neighbourhood SES	Barriers to housing and services decile	Continuous	Y
Neighbourhood SES	Living environment decile	Continuous	Y
Attitude child education	Child plays musical instrument	Binary	Y
Attitude child education	Takes part in extracurricular activities	Binary	Y
Attitude child education	Most common method of praise given to child	Binary	Y
Attitude child education	Designated homework space	Binary	N
Attitude child education	Ideal family day out scoring educational	Binary	Y
Attitude child education	Childq1 attitude school work	Polytomous	Y
Attitude child education	Childq6 attitude school	Polytomous	Y
Attitude child education	Importance of profession to caregiver	Polytomous	N
Attitude child education	Importance of caregiver own educational level	Polytomous	N
Attitude child education	Worry about child school performance	Polytomous	Y
Attitude child education	Childs attitude school code	Polytomous	Y
Attitude child education	Help with homework	Polytomous	Y
Attitude child education	Parent attitude homework	Polytomous	Y
Attitude child education	School code	Polytomous	N

Child health	Mother smoked in pregnancy	Binary	Y
Child health	Child eats breakfast everyday	Binary	Y
Child health	Ideal family day out scoring active	Binary	N
Child health	Childq2 attitude appearance	Polytomous	N
Child health	Childq7 attitude overall wellbeing	Polytomous	N
Child health	Portions of fruit and veg per day	Polytomous	Y
Child health	Birthterm	Polytomous	Y
Child health	Days off school	Polytomous	Y
Child health	Current child health	Polytomous	Y
Child health	Child health throughout life	Polytomous	Y
Child health	Total unhealthy portions food	Continuous	Y
Child health	Child birth weight kg	Continuous	Y
Child health	Number of hours child sleeps	Continuous	Y
Child health	Days a week child plays outside	Continuous	Y
Child health	Paternal age at birth	Continuous	Y
Child health	Maternal age at birth	Continuous	Y
Primary caregiver wellbeing	Caregiver finds life stressful on average	Binary	Y

Primary caregiver wellbeing	Caregiver gets enough support	Polytomous	Y
Primary caregiver wellbeing	Primary alcohol consumption	Polytomous	N
Primary caregiver wellbeing	Caregiver participation in religion	Polytomous	N
Primary caregiver wellbeing	Caregiver conflict with partner	Polytomous	N
Primary caregiver wellbeing	Age primary caregiver	Continuous	N
Primary caregiver wellbeing	Number of alcohol units	Continuous	N
Primary caregiver wellbeing	Holmes Rahe stress child's life	Continuous	Y
Primary caregiver wellbeing	Holmes Rahe stress last year	Continuous	Y
Time with family and friends	Child regularly plays with other children	Binary	Y
Time with family and friends	How often child sees grandparents	Polytomous	Y
Time with family and friends	How often child sees aunts uncles	Polytomous	Y
Time with family and friends	How often child sees cousins	Polytomous	Y
Time with family and friends	How often child sees family friends	Polytomous	Y
Time with family and friends	Primary evening work code	Polytomous	Y

Time with family and friends	Caregiver and child activity together total	Continuous	Y
Time with family and friends	Days a week eat dinner as family	Continuous	Y
Opinion family relationships	Childq3 attitude parents	Polytomous	Y
Opinion family relationships	Childq4 attitude siblings	Polytomous	Y
Opinion family relationships	Childq5 attitude friends	Polytomous	Y
Opinion family relationships	Caregiver opinion how well they listen to child	Polytomous	Y
Opinion family relationships	Converse with child about their day	Polytomous	Y
Opinion family relationships	Caregiver opinion of relationship with family	Polytomous	Y
Opinion family relationships	Caregiver opinion of relationship with child	Polytomous	Y
Childcare	Care by childcare centre or after school club	Binary	Y
Childcare	Age child started regular childcare	Continuous	Y
Childcare	Care by other relative 16years or over	Binary	N
Time with family and friends	Child has set dinner time	Binary	N
Time with family and friends	Child sits at table for dinner	Binary	N

Relationships additional	Primary caregiver gender	Binary	Y
Relationships additional	Child shares bedroom	Binary	Y
Relationships additional	Partner evening work code	Polytomous	Y
Relationships additional	Child sibling position	Continuous	Y
Technology use	Child has mobile phone	Binary	Y
Technology use	Child has TV or computer in bedroom	Binary	Y
Technology use	Access to games consoles	Binary	Y
Technology use	Number household computer tablet laptop	Polytomous	Y
Technology use	Child internet usage time	Polytomous	Y
Technology use	Number of televisions in house	Polytomous	Y
Technology use	Hours of TV per day	Polytomous	Y
Technology use	Uses internet social media and chat	Polytomous	Y
Technology use	Uses internet for fun	Polytomous	Y
Technology use	Uses internet for education	Polytomous	Y
Discipline	Discipline by explanation	Polytomous	Y
Discipline	Discipline by shouting	Polytomous	Y
Discipline	Discipline by grounding	Polytomous	Y
Discipline	Discipline by hitting smacking	Polytomous	Y

Discipline	Discipline by limiting technology time	Polytomous	Y
Discipline	Discipline by timeout	Polytomous	Y
Discipline	Let child cause trouble	Polytomous	Y
Other language use	Caregiver other languages at home	Binary	Y
Other language use	Child other languages at home	Binary	Y
Other language use	Level of English at home	Polytomous	Y
Reading at home	Daily newspaper	Binary	Y
Reading at home	Books in the home	Polytomous	Y
Reading at home	Caregiver attitude child reading	Polytomous	Y
Reading at home	Reading together outloud	Polytomous	Y
Reading at home	Minutes per day child read to	Polytomous	Y
Reading at home	Minutes per day child reads alone	Polytomous	Y
Reading at home	Parental interest in own reading	Polytomous	Y
Reading at home	Child interest in reading	Polytomous	Y
Reading at home	Use of local library	Continuous	Y
Attitude to neighbourhood	Sense of neighbourhood belonging	Polytomous	Y
Attitude to neighbourhood	Friends in neighbourhood	Polytomous	Y

Attitude to neighbourhood	Advice in neighbourhood	Polytomous	Y
Attitude to neighbourhood	Neighbour relationships	Polytomous	Y
Attitude to neighbourhood	Willing to work with those in neighbourhood	Polytomous	Y
Attitude to neighbourhood	Similar to others in neighbourhood	Polytomous	Y
Attitude to neighbourhood	Converse with others in neighbourhood	Polytomous	Y
Rules and chores	Rules about what child can watch	Binary	Y
Rules and chores	Child not allowed to go out unsupervised	Binary	N
Rules and chores	Child makes bed	Binary	Y
Rules and chores	Child cleans room	Binary	Y
Rules and chores	Child bathes themselves	Binary	N
Rules and chores	Child has routine chores	Binary	Y
Rules and chores	Monitored or supervised access computer	Binary	Y
Rules and chores	Always know where child is going out	Polytomous	N

B.1 Detailed description of the process used to form the environmental domains based on the factor analyses.

The questionnaire items were divided into the following domains: attitude to child education, attitude to neighbourhood, child health, discipline, rules and chores, neighbourhood SES, other language use and reading, parent health and wellbeing, social resources, family relationships and use of technology. These were submitted to the quasi-data driven factor analysis method explained in section 2.2.3.4 and the results are detailed here.

Attitude to neighbourhood, neighbourhood SES, and use of technology each had a clean factor structure in which only the first factor had an eigenvalue greater than 1. The parent health factor was almost entirely dominated by measures of parent wellbeing and stress and so, in order to obtain a cleaner factor, the items with low loadings (*Age primary caregiver, Number of alcohol units, Primary alcohol consumption and Caregiver participation in religion*) were removed and the domain was renamed 'parent wellbeing'. Note that the loadings were multiplied by -1 to ensure that a high score on this measure relates to high wellbeing and low stress. In addition, *conflict with partner* was removed as there were a number of families with single parents.

For the remaining domains, more than one factor had an eigenvalue greater than 1. If the items separated into a factor structure (obtained using Promax rotation) that related to clear sub-domains, the items were further divided between these and factor analyses were repeated on each set. Thus, discipline and rules and chores, use of other languages and reading at home, and family relationships were divided into discipline, rules and chores, other language use, reading at home, opinion of family relationships, time with family and friends and childcare respectively. The social resources domain separated into two factors: the personal Resource Generator-UK subscale and a factor with all of the other subscales and so the personal social resources measure was used as a separate measure (named primary caregiver personal skills) and the factor analysis was repeated on the remaining subscales.

Furthermore, some of the items were excluded if they had a very low loading or were thought to be relatively poor measures of the domain until only one factor had an eigenvalue >1. *Importance of profession to caregiver* and *Importance of caregiver own educational level* were excluded from the attitude to child education domain as they dominated the factor structure yet aren't necessarily measures of child education. In addition, *designated homework space* was removed as it had a small negative loading

and it is not clear how this might relate to positive child attitude to education. Finally, *child school* was removed as it did not directly relate to attitude to child education. *Child attitude to appearance and Child attitude to overall wellbeing* were removed from the child health domain as, again, they dominated the factors but are arguably not the cleanest measures of child health in comparison to the other questions. *Ideal family day out scoring educational* ,a binary measure of whether the primary caregiver's ideal family day out included educational activities, was removed as it had a low loading and removal resulted in only one factor with an eigenvalue >1. *Child has set dinner time, Child sits at table for dinner* were excluded from the time with family and friends domain as they are likely to be relatively poor measures yet the high correlation between them resulted in high negative loadings. *Care by other relative 16 years* or over was removed from the childcare domain as the loading was very low and it was the only question that survived the skewness criterion from the set of questions about childcare within the family. Finally, *Child bathes themselves*, *child not allowed to go out unsupervised,* and *Always know where child is when they're outside* were all excluded from rules and chores as they are likely to be highly influenced by child age.

After removing these items, each of these domains had only one factor with eigenvalue >1. The majority of the factor analysis scree plots and parallel analyses also supported the decision to extract a single factor for each domain. Where these secondary conditions were not met (e.g. there was no discernible elbow in the scree plot), we checked that the additional variance explained by two factors in comparison to the first factor was minimal.

B.2 The item correlations and factor analysis loadings for each environmental domain
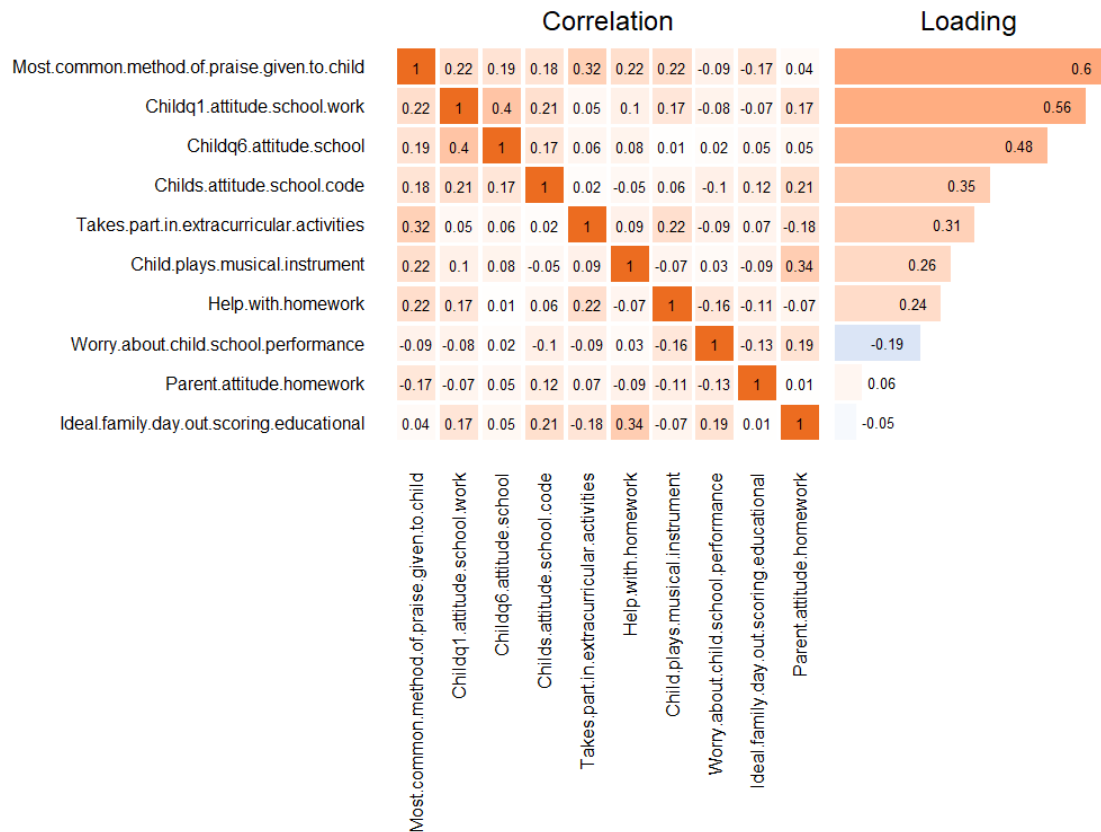

Figure 72 Attitude to child education correlation table and factor loadings
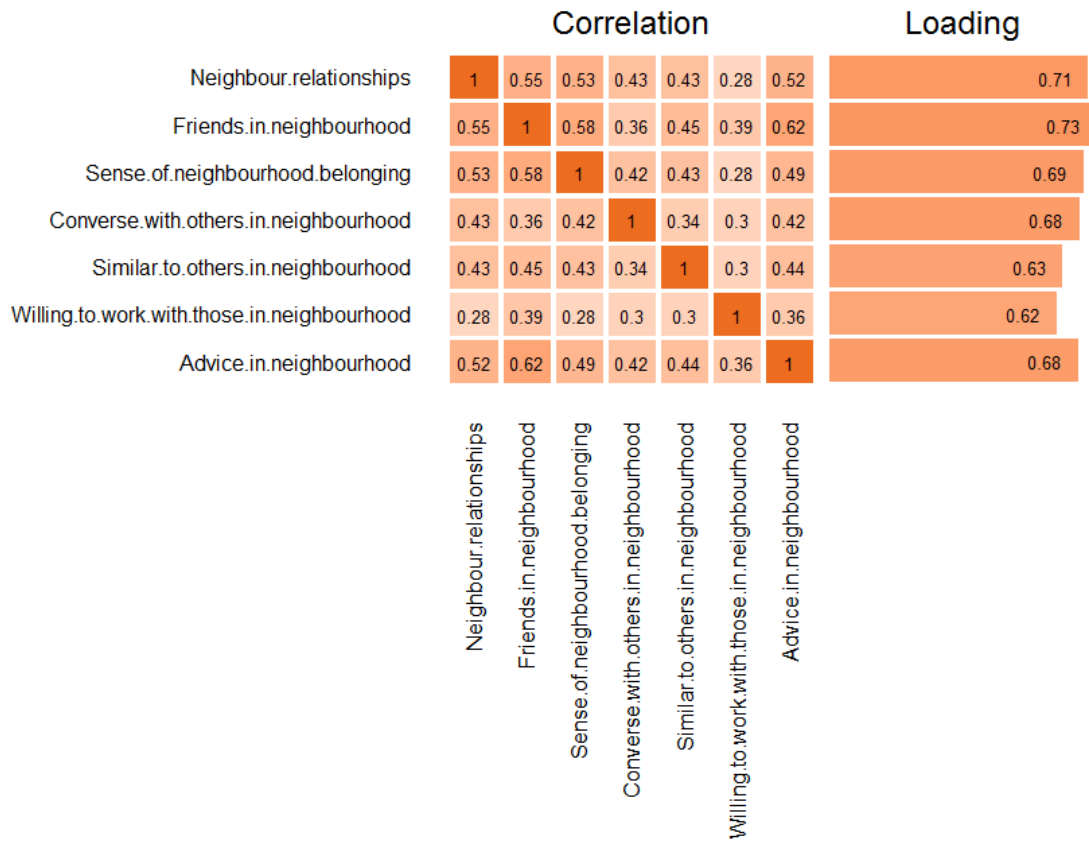
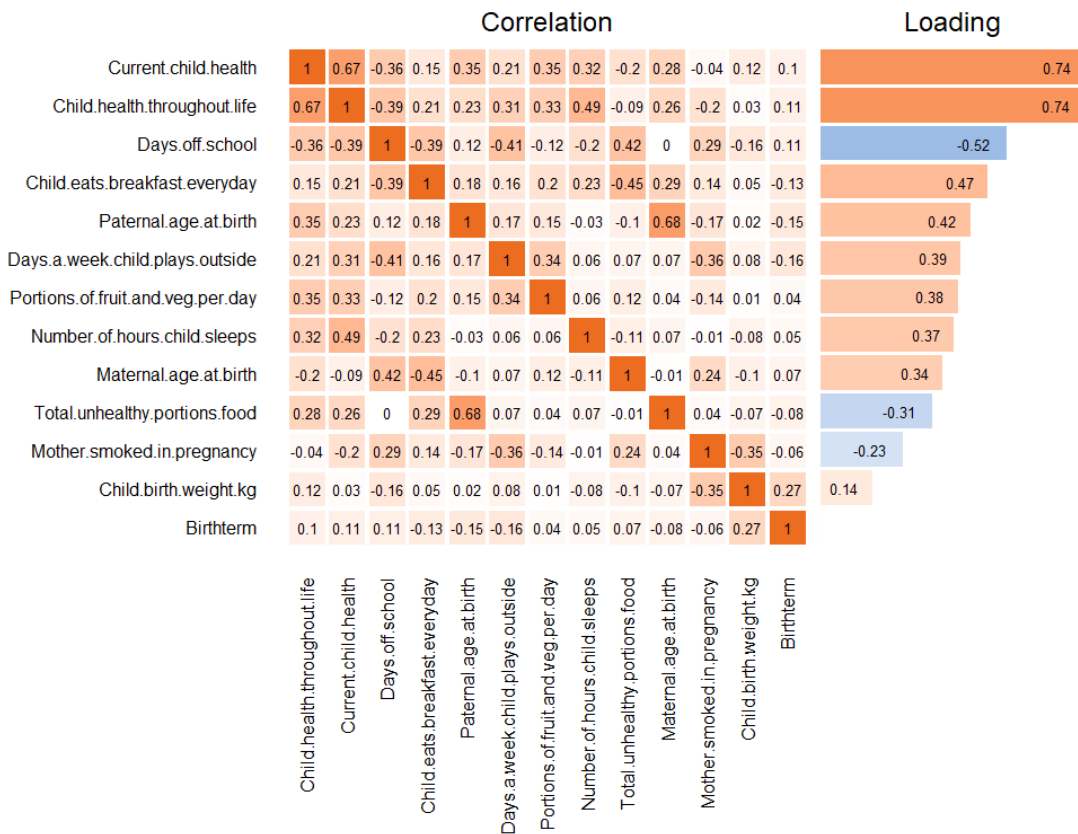**Figure 73** Attitude to neighbourhood correlation table and factor loadings

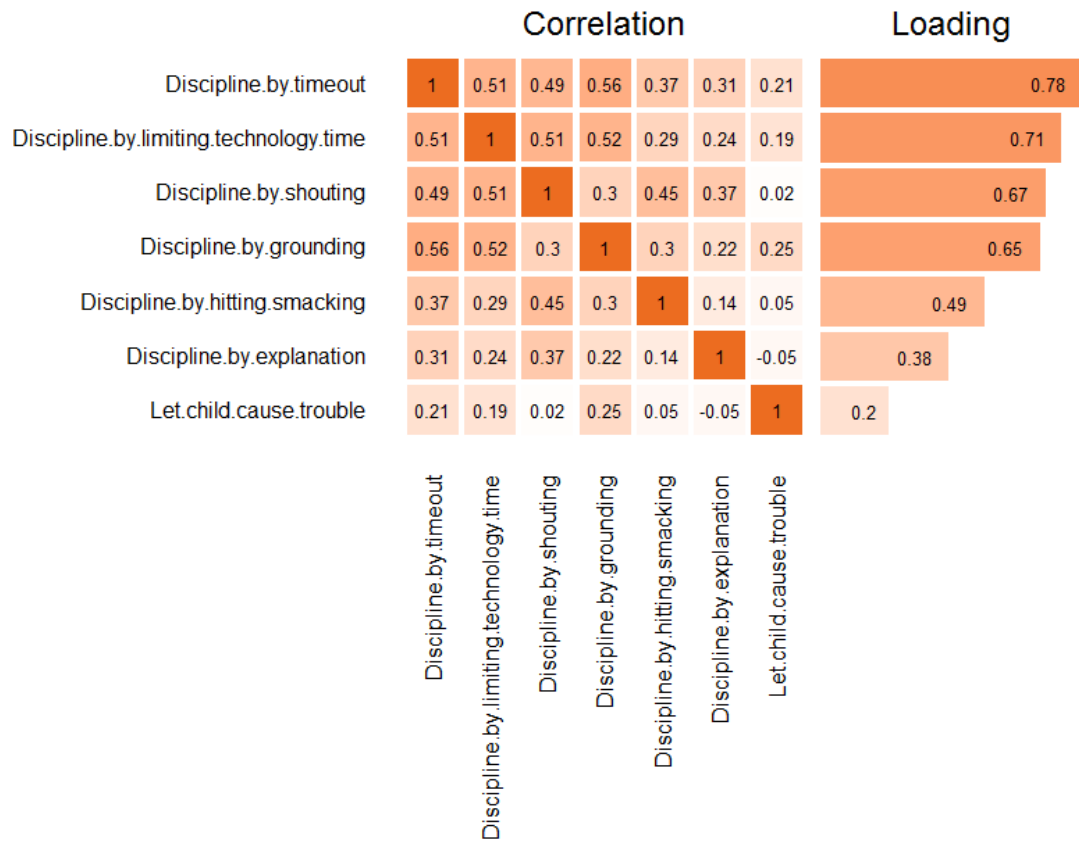**Figure 74** Child health correlation table and factor loadings

Correlation   Loading

	Discipline.by.timeout	Discipline.by.limiting.technology.time	Discipline.by.shouting	Discipline.by.grounding	Discipline.by.hitting.smacking	Discipline.by.explanation	Let.child.cause.trouble	Loading
Discipline.by.timeout	1	0.51	0.49	0.56	0.37	0.31	0.21	0.78
Discipline.by.limiting.technology.time	0.51	1	0.51	0.52	0.29	0.24	0.19	0.71
Discipline.by.shouting	0.49	0.51	1	0.3	0.45	0.37	0.02	0.67
Discipline.by.grounding	0.56	0.52	0.3	1	0.3	0.22	0.25	0.65
Discipline.by.hitting.smacking	0.37	0.29	0.45	0.3	1	0.14	0.05	0.49
Discipline.by.explanation	0.31	0.24	0.37	0.22	0.14	1	-0.05	0.38
Let.child.cause.trouble	0.21	0.19	0.02	0.25	0.05	-0.05	1	0.2

Figure 75 Discipline correlation table and factor loadings

Correlation   Loading

	Child.has.routine.chores	Child.cleans.room	Monitored.or.supervised.access.computer	Child.makes.bed	Rules.about.what.child.can.watch	Loading
Child.has.routine.chores	1	0.43	0.4	0.5	0.3	0.8
Child.cleans.room	0.43	1	0.37	0.45	0.23	0.68
Monitored.or.supervised.access.computer	0.4	0.37	1	0	0.31	0.47
Child.makes.bed	0.5	0.45	0	1	-0.15	0.47
Rules.about.what.child.can.watch	0.3	0.23	0.31	-0.15	1	0.31
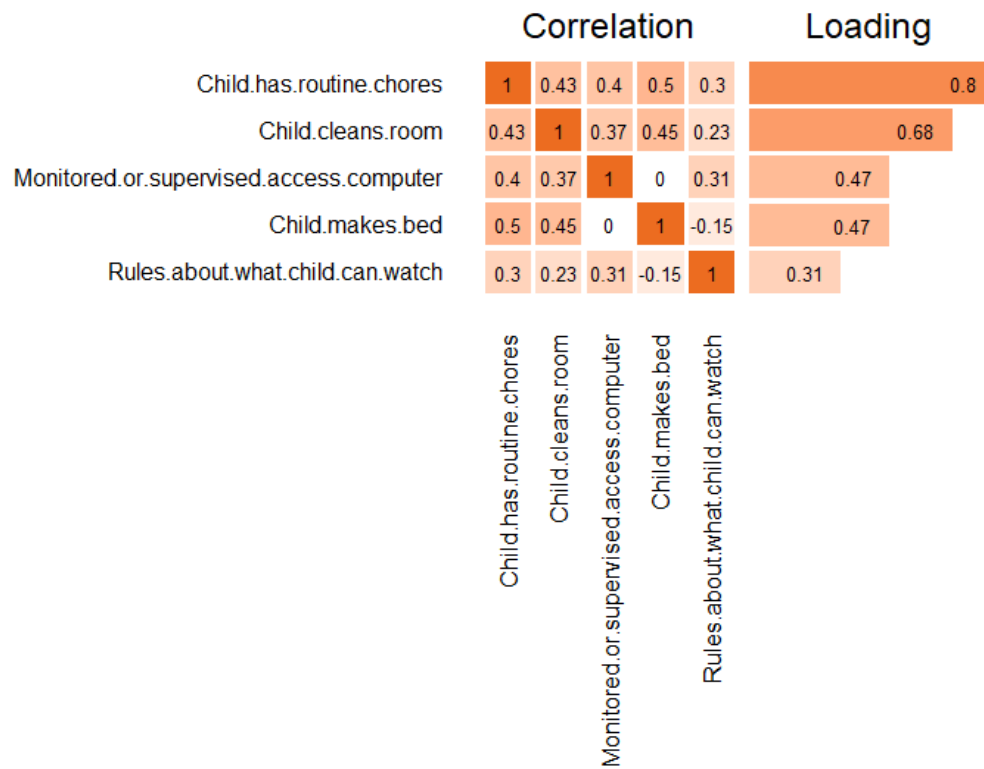
Figure 76 Rules and chores correlation table and factor loadings

Figure 77 Other language use correlation table and factor loadings

Figure 78 Reading at home correlation table and factor loadings

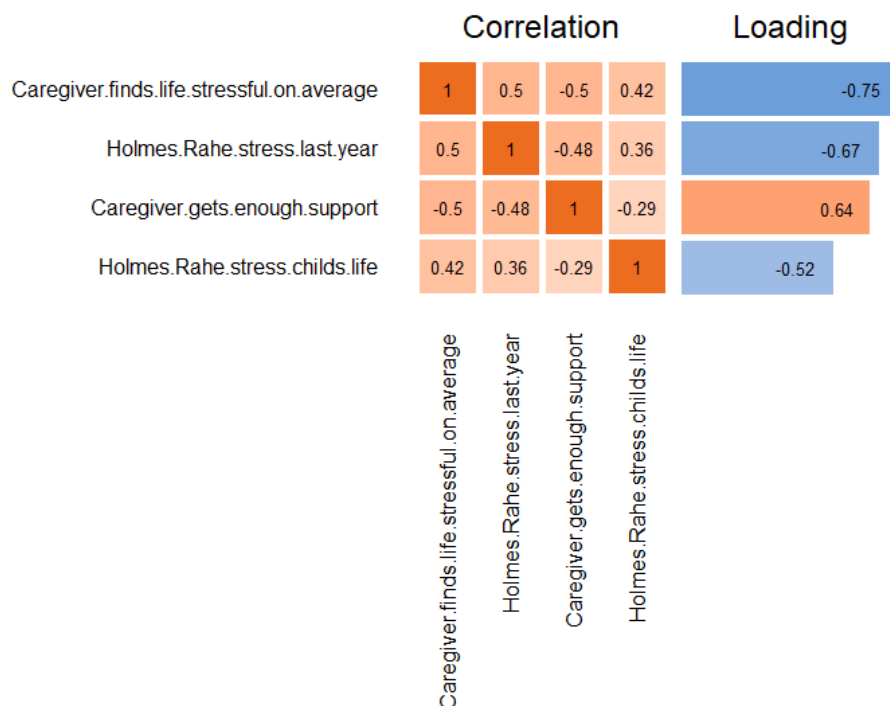**Figure 79** Neighbourhood SES correlation table and factor loadings

Figure 80 Primary caregiver wellbeing correlation table and factor loadings

Figure 81 Acquaintance skills (Resource Generator) correlation table and factor loadings

Figure 82 Opinion of family relationships correlation table and factor loadings
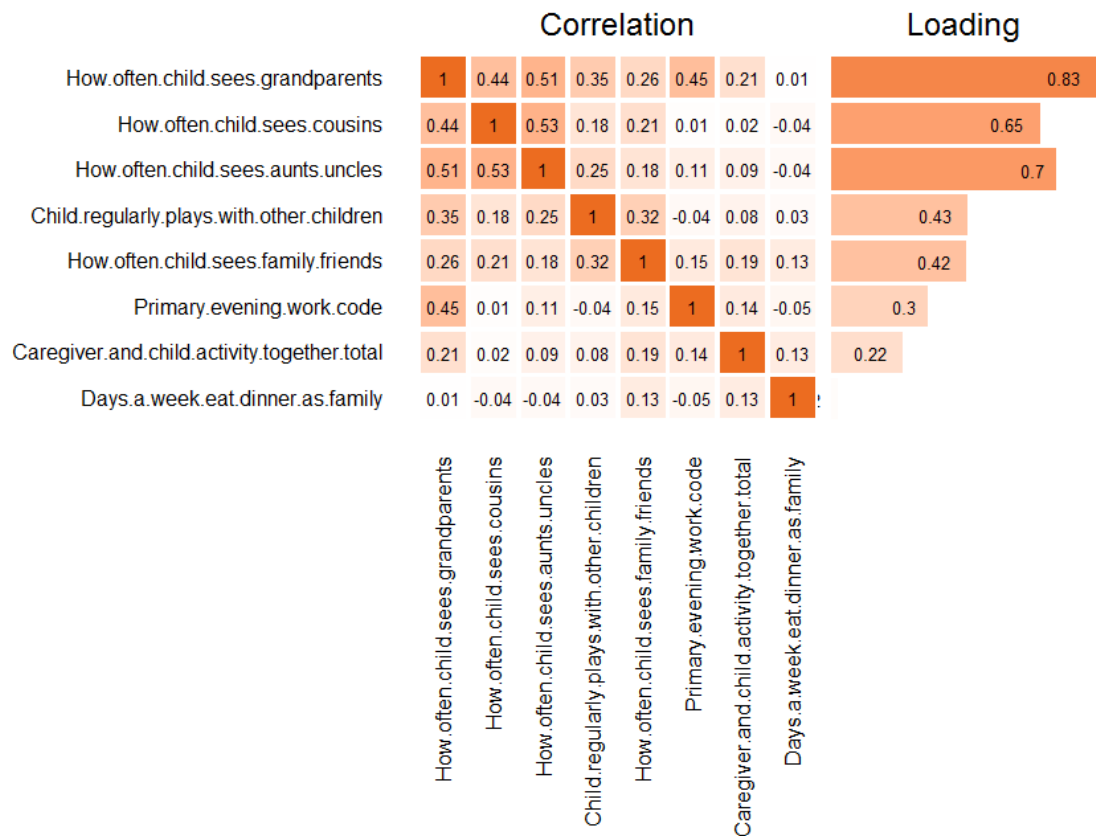
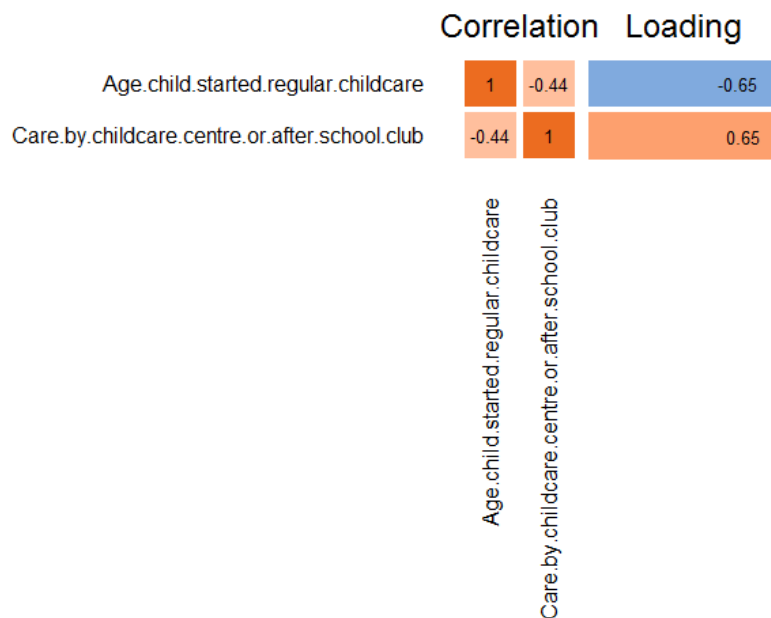**Figure 83** Time with family and friends correlation table and factor loadings

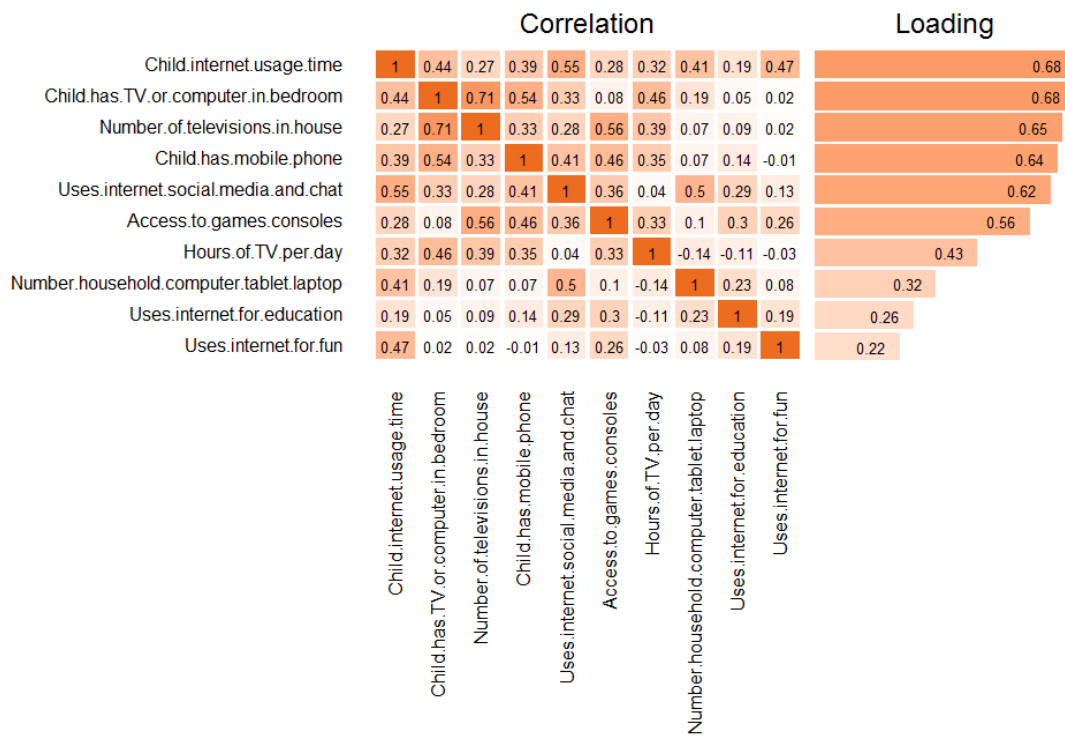**Figure 84** Childcare correlation table and factor loadings

Figure 85 Technology use correlation table and factor loadings

APPENDIX C: THE ACADEMIC, BEHAVIOUR AND COGNITIVE SUBTESTS

Table 10 The child developmental measures and their corresponding subtests.

Outcome	Subtest	Tool
Working Memory	WM DR std	AWMA
Working Memory	WM DM std	AWMA
Working Memory	WM MX std	AWMA
Working Memory	WM BD std	AWMA
Vocab IQ	Vocab T score	WASI
Matrix IQ	Matrix T score	WASI
Phonological processing	PhAB Naming speed std	PhAB
WJ Reading	WJ Reading	WJ
WJ Maths	WJ Maths	WJ
BRIEF	BRIEF Inhibit	BRIEF
BRIEF	BRIEF Shift	BRIEF
BRIEF	BRIEF Emotional control	BRIEF
BRIEF	BRIEF Initiate	BRIEF
BRIEF	BRIEF WM	BRIEF
BRIEF	BRIEF Plan organize	BRIEF
BRIEF	BRIEF Org of materials	BRIEF
BRIEF	BRIEF Monitor	BRIEF
SDQ	SDQ Emotion problems	SDQ
SDQ	SDQ Conduct problems	SDQ
SDQ	SDQ Hyperactivity	SDQ
SDQ	SDQ Peer problems	SDQ
SDQ	SDQ Prosocial	SDQ

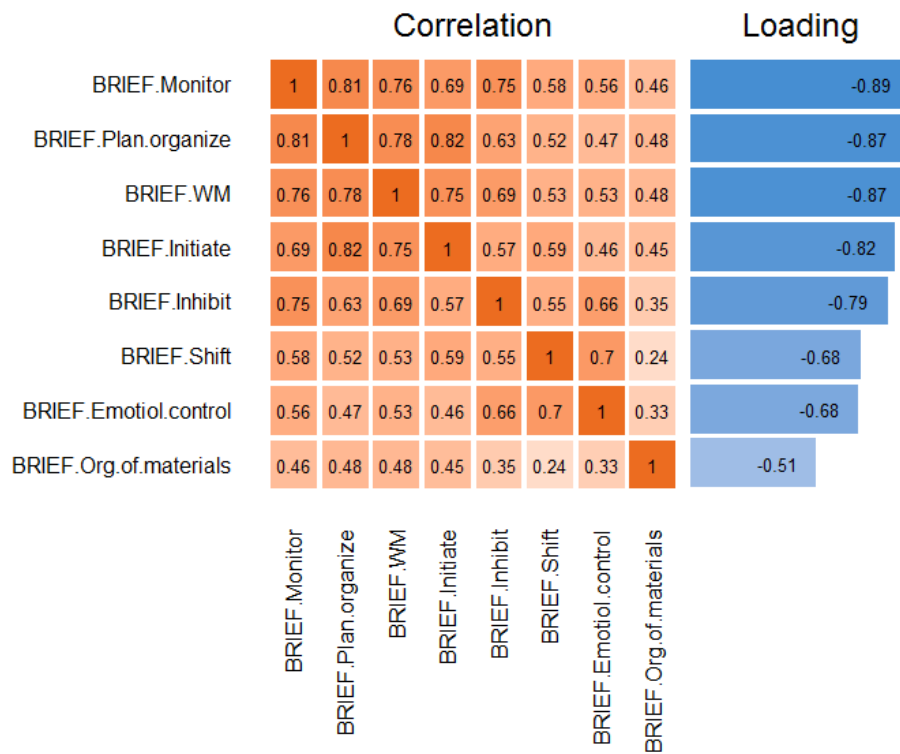## C.1 The item correlations and factor analysis loadings for each child development measure

Figure 86 BRIEF correlation table and factor loadings
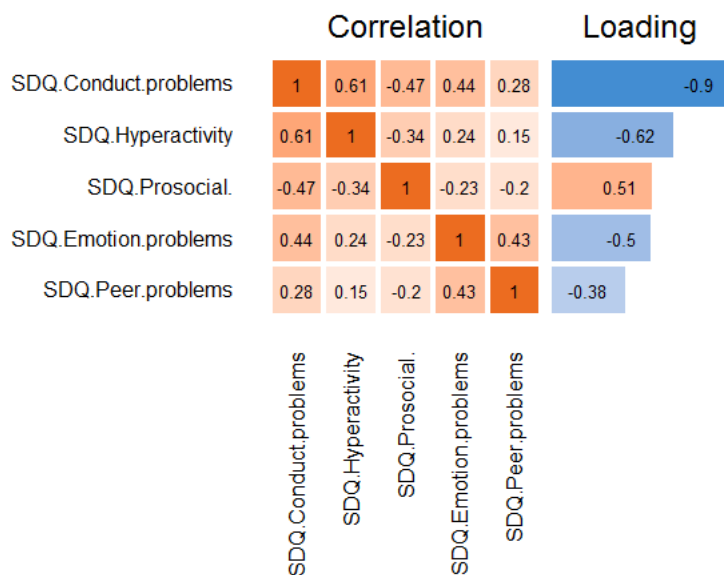
Figure 87 SDQ correlation table and factor loadings

Figure 88 Working memory correlation table and factor loadings

APPENDIX D: LASSO PREDICTION OF CHILD OUTCOMES USING AIC
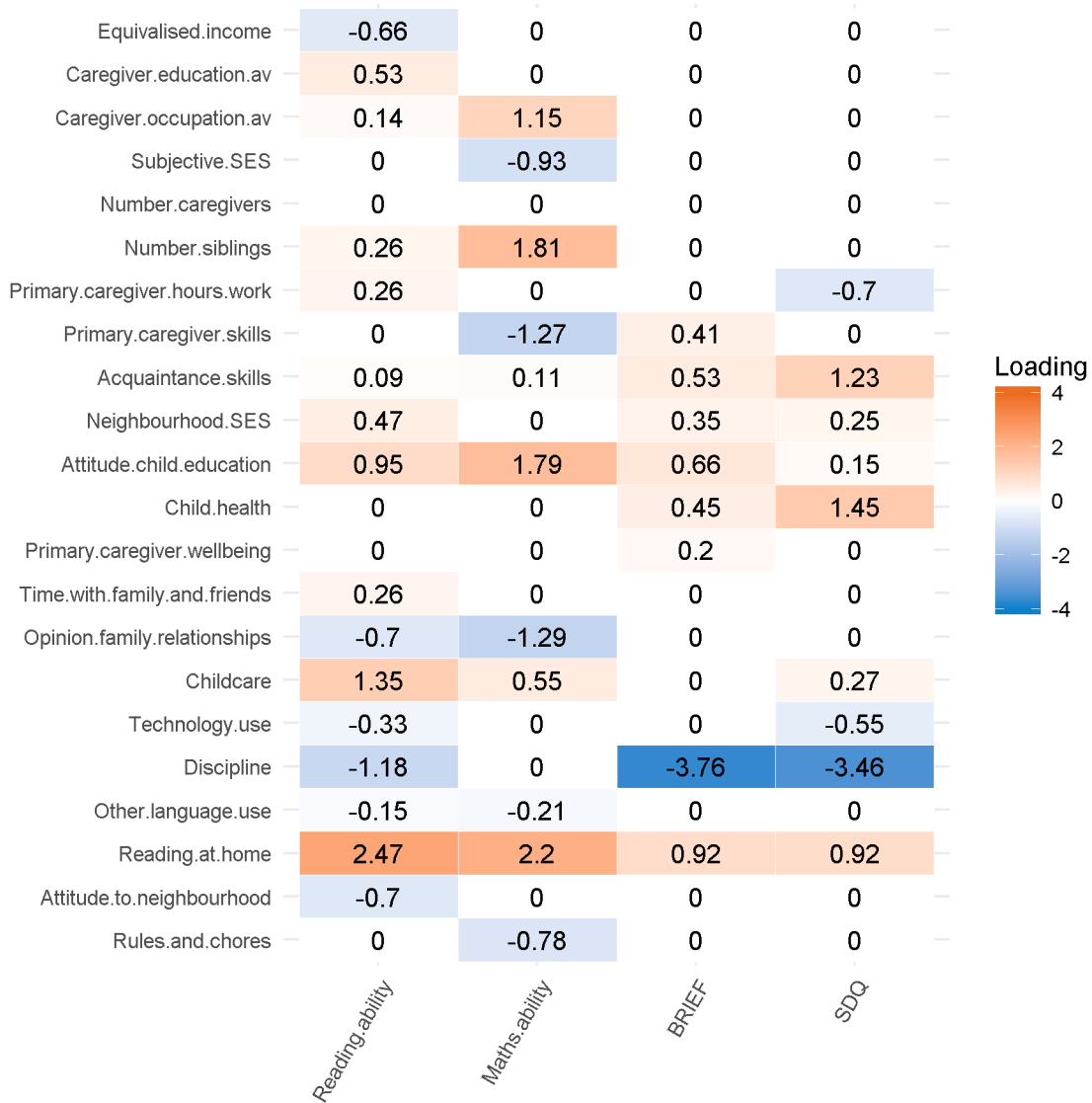

Figure 89 The selected LASSO regression coefficients for the model with the lowest AIC value

APPENDIX E: COGNITION AND BEHAVIOR SUBTEST CORRELATIONS

	WM.DR.std	WM.DM.std	WM.MX.std	WM.BD.std	Vocab.T.score	Matrix.T.score	PhAB.Naming.speed.std	BRIEF.Inhibit	BRIEF.Shift	BRIEF.Emotiol.control	BRIEF.Initiate	BRIEF.WM	BRIEF.Plan.organize	BRIEF.Org.of.materials	BRIEF.Monitor	SDQ.Emotion.problems	SDQ.Conduct.problems	SDQ.Hyperactivity	SDQ.Peer.problems	SDQ.Prosocial.
WM.DR.std	1	0.2	0.23	0.43	0.29	0.42	0.18	0.08	0.01	0.08	0.14	0.15	0.04	-0.21	0.02	-0.05	0.13	0.07	-0.08	-0.1
WM.DM.std	0.2	1	0.39	0.45	0.15	0.44	0.17	0.08	0.12	0.1	0.13	0.22	0.16	0.07	0.21	0.11	0.12	0.16	0.06	-0.06
WM.MX.std	0.23	0.39	1	0.31	0.26	0.45	0.1	-0.06	0.12	-0.02	0.17	0.21	0.19	0.03	0.25	0.07	-0.07	0.04	0.01	-0.21
WM.BD.std	0.43	0.45	0.31	1	0.39	0.37	0.35	0.04	0.05	0.03	0.15	0.11	0.12	0	0.12	0.03	0.13	0.09	0.08	-0.15
Vocab.T.score	0.29	0.15	0.26	0.39	1	0.51	0.34	0.13	0.15	0.02	0.34	0.26	0.24	0.08	0.26	0.11	0.04	0.2	0.17	0.03
Matrix.T.score	0.42	0.44	0.45	0.37	0.51	1	0.3	0.24	0.2	0.14	0.29	0.39	0.3	0.03	0.42	0.13	0.16	0.22	0.13	0.03
PhAB.Naming.speed.std	0.18	0.17	0.1	0.35	0.34	0.3	1	0.07	-0.04	-0.06	0.1	0.14	0.14	0.04	0.1	-0.04	0	0.04	0.13	-0.1
BRIEF.Inhibit	0.08	0.08	-0.06	0.04	0.13	0.24	0.07	1	0.54	0.66	0.57	0.69	0.63	0.35	0.75	0.2	0.67	0.78	0.24	0.42
BRIEF.Shift	0.01	0.12	0.12	0.05	0.15	0.2	-0.04	0.54	1	0.7	0.59	0.53	0.52	0.24	0.58	0.5	0.52	0.53	0.41	0.35
BRIEF.Emotiol.control	0.08	0.1	-0.02	0.03	0.02	0.14	-0.06	0.66	0.7	1	0.46	0.53	0.47	0.33	0.56	0.55	0.63	0.64	0.26	0.33
BRIEF.Initiate	0.14	0.13	0.17	0.15	0.34	0.29	0.1	0.57	0.59	0.46	1	0.75	0.82	0.45	0.69	0.3	0.42	0.65	0.3	0.31
BRIEF.WM	0.15	0.22	0.21	0.11	0.26	0.39	0.14	0.69	0.53	0.53	0.75	1	0.78	0.48	0.76	0.23	0.45	0.74	0.24	0.2
BRIEF.Plan.organize	0.04	0.16	0.19	0.12	0.24	0.3	0.14	0.63	0.52	0.47	0.82	0.78	1	0.48	0.81	0.15	0.48	0.74	0.27	0.24
BRIEF.Org.of.materials	-0.21	0.07	0.03	0	0.08	0.03	0.04	0.35	0.24	0.33	0.45	0.48	0.48	1	0.46	0.26	0.27	0.42	0.11	0.16
BRIEF.Monitor	0.02	0.21	0.25	0.12	0.26	0.42	0.1	0.75	0.58	0.56	0.69	0.76	0.81	0.46	1	0.28	0.53	0.74	0.28	0.29
SDQ.Emotion.problems	-0.05	0.11	0.07	0.03	0.11	0.13	-0.04	0.2	0.5	0.55	0.3	0.23	0.15	0.26	0.28	1	0.42	0.24	0.42	0.22
SDQ.Conduct.problems	0.13	0.12	-0.07	0.13	0.04	0.16	0	0.67	0.52	0.63	0.42	0.45	0.48	0.27	0.53	0.42	1	0.62	0.27	0.46
SDQ.Hyperactivity	0.07	0.16	0.04	0.09	0.2	0.22	0.04	0.78	0.53	0.64	0.65	0.74	0.74	0.42	0.74	0.24	0.62	1	0.14	0.34
SDQ.Peer.problems	-0.08	0.06	0.01	0.08	0.17	0.13	0.13	0.24	0.41	0.26	0.3	0.24	0.27	0.11	0.28	0.42	0.27	0.14	1	0.2
SDQ.Prosocial.	-0.1	-0.06	-0.21	-0.15	0.03	0.03	-0.1	0.42	0.35	0.33	0.31	0.2	0.24	0.16	0.29	0.22	0.46	0.34	0.2	1

Figure 90 Correlation between each cognitive subtest and the BRIEF and SDQ subscales.